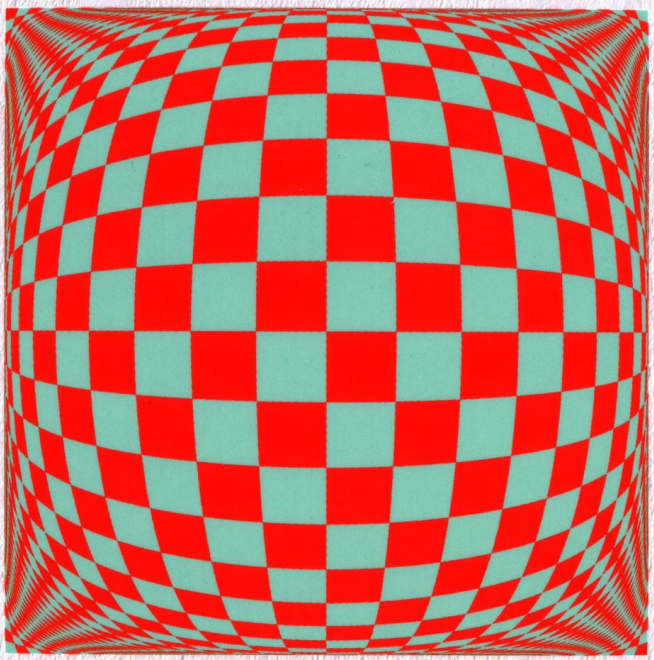
HANNAH FRY

HELLO, WORLD!

Revoluția informatică și viitorul omenirii



CORINT FUTURE

„Cu o simplitate de-a dreptul înviorătoare, Hannah Fry reușește să explice unele dintre cele mai spinoase probleme ale lumii informatice moderne, de la deciziile luate de inteligența artificială, rețelele neurale și învățarea automată, la scandalul stârnit de implicarea firmei britanice Cambridge Analytica în alegerile americane din 2016, neajunsurile automobilelor autonome și multe altele. Această carte ne demonstrează tuturor de ce este imperios necesar să existe oameni de știință excelenți precum Hannah Fry, care să se dedice, în același timp, și scrisului."

KATY GUEST, *Guardian*

HANNAH FRY

HELLO, WORLD!

CORINT FUTURE

Colecţie coordonată de
ADA ROSETI

HANNAH FRY

HELLO, WORLD!

Revoluția informatică și viitorul omenirii

Traducere din limba engleză
ADINA IHORA

Cuvânt-înainte
CIPRIAN I. TOMOIOAGĂ

CORINT BOOKS
—2019—

Lectură de specialitate: Ciprian I. Tomoioagă
Redactare: Alina Sârbu, Alina Bâltâc
Glosar și note: Alina Sârbu, Ciprian I. Tomoioagă
Tehnoredactare: Zamfir Adriana Nicoleta PFA
Design copertă: Dan Mihalache
Ilustrația copertei: © Shutterstock
Ilustrații interior: p. 27, *Mașina–câine,* fotografie reprodusă cu permisiunea lui Danilo Vasconcellos Vargas, Universitatea Kyushu, Fukuoka, Japonia; p. 122, *Gorilă-n piept,* fotografie reprodusă cu permisiunea lui Trafton Drew, Universitatea din Utah, Salt Lake City, SUA; p. 208, *Imaginile lui Steve Talley* © Steve Talley (stânga) și FBI; p. 211, *Neil Douglas și sosia lui,* fotografie reprodusă cu permisiunea lui Neil Douglas; p. 216, *Ochelarii „Țestoasă",* fotografie reprodusă cu permisiunea lui Mahmood Sharif, Universitatea Carnegie Mellon, Pittsburgh, SUA; p. 216, Georges Biard, *Milla Jovovich la Festivalul de Film de la Cannes,* fotografie reprodusă cu permisiunea autorului.

HANNAH FRY
Hello World: Being Human in the Age of Algorithms
Copyright © Hannah Fry Limited 2018
All rights reserved.

Toate drepturile asupra ediției în limba română aparțin
EDITURII CORINT BOOKS.

ISBN: 978-606-793-598-1

Descrierea CIP a Bibliotecii Naționale a României
FRY, HANNAH
Hello, world! : revoluția informatică și viitorul omenirii / Hannah Fry ; trad. din lb. engleză: Adina Ihora ; cuv.-înainte și lectură de specialitate: Ciprian I. Tomoioagă. - București : Corint Books, 2019
Index
ISBN 978-606-793-598-1

I. Ihora, Adina (trad.)
II. Tomoioagă, Ciprian I. (pref.)

004

CUVÂNT-ÎNAINTE

Vi s-a întâmplat, uneori, să ajungeți acasă și să simțiți că aveți chef de stat pe Facebook sau pe Instagram, fără să știți ce anume ați vrea să vedeți acolo? Sau să deschideți o asemenea aplicație, doar ca să o închideți în mai puțin de două minute? Oare cum am dobândit astfel de automatisme?

În capitolul al doilea al acestei minunate introduceri în lumea algoritmilor pentru ne-programatori*, Hannah Fry oferă mai mult decât răspunsuri la amintita întrebare. Folosind explicații simple, însoțite de exemple concise, ea reușește să scoată la iveală cele mai grave probleme pe care internetul și universul on-line ni le-au creat recent. De la faimoasele scandaluri legate de confidențialitate ale companiei Facebook și până la practicile obscure ale magazinelor (on-line și off-line), aceste istorioare ne ajută să înțelegem dinamicile complexe care sunt folosite azi pentru a ne face să ne cheltuim cea mai prețioasă monedă: atenția.

Citind sus-amintitul capitol, am trecut printr-un amalgam de stări emoționale. Întâi, am fost bucuros că autoarea aduce

* Deși cartea de față oferă detalii despre algoritmi, aceasta nu este legată în niciun fel de opera foarte apreciată de programatori *Introduction to Algorithms* („Algoritmii. O introducere", MIT Press & McGraw Hill, Cambridge, Massachussets, 1990) de Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest și Clifford Stein, care servește drept manual de programare la cea mai prestigioasă universitate politehnică din lume, MIT (Massachussetts Institute of Technology).

în atenția publicului larg probleme despre care eu, ca informatician, sunt conștient, dar nu am puterea de a le mediatiza. Apoi, am fost uluit să aflu că unele practici, precum utilizarea cardului de fidelitate la un magazin, au apărut cu mult înaintea internetului așa cum îl știm azi (1993). Am fost entuziasmat și speriat, în același timp, când am văzut că algoritmul de recomandare al magazinului Target a depistat sarcina unei adolescente înaintea familiei acesteia, și că asta se întâmpla în anul 2002! Entuziasmul s-a pierdut și a rămas doar îngrijorarea când am realizat că, la câte reacții de *Apreciere*, *Râs*, *Plâns* și *Furie* am lăsat pe Facebook, platforma de socializare a ajuns să mă cunoască mai bine decât propriii părinți! Iar îngrijorarea s-a transformat apoi în furie, aflând că analiștii folosesc modelul meu psihologic pentru a-mi influența opiniile politice, și m-a cuprins deznădejdea când am aflat că este imposibil, în momentul de față, „să resetezi" ce știe internetul despre tine, așa cum a demonstrat-o cazul unei mame care a pierdut sarcina mult dorită, dar nu a reușit să scape de reclamele destinate gravidelor, care o hăituiau peste tot online. În final, mi-am recăpătat speranța că nu e totul pierdut și încă putem lua măsuri, cum este legea cu privire la protecția datelor în Europa. Trebuie doar să nu stăm cu mâinile în sân și să fim conștienți că datele noastre personale sunt, în ziua de azi, mai prețioase decât petrolul. Cartea aceasta ne ajută să înțelegem că suntem, fiecare în parte și toți împreună, foarte, foarte bogați și puternici. Depinde de noi, doar să le cerem parlamentarilor legi rezonabile, care să ne protejeze, fără să ne îngrădească.

În ciuda sentimentelor diverse prin care ne poate purta, *Hello, world!* nu se vrea a fi un articol à la BuzzFeed*. Dimpotrivă, am apreciat mult numeroasele anecdote care te ajută să înțelegi

* Site de media și divertisment specializat în articole virale, binecunoscut pentru practicile sale controversate de a atrage cititorii, bazate pe crearea de conținut emoționant, dar uneori incorect.

și să te raportezi și afectiv la chestiunile prezentate, întrucât nu cred că e cineva expert în toate cele șapte domenii abordate. În plus, este evident spiritul de cercetătoare al autoarei din faptul că absolut toate anecdotele sunt cazuri reale, cu conținut verificabil și cu surse citate în notele de la final. Ba, uneori, autoarea face trimitere chiar la filmulețe postate pe YouTube, pe care le-am verificat, ca un tocilar pasionat, și pot mărturisi că merită efortul de a transcrie linkul în navigator!* Trebuie observat, însă, că situațiile prezentate sunt mai mult decât niște simple povești de seară sau subiecte de discuție la petreceri. Sunt convins că autoarea le-a ales cu grijă pentru a putea aborda răspunsuri diverse la întrebări dificile, pe care, uneori, ne este frică să le punem. Spre exemplu, în dilema vagonului scăpat de sub control**, nu există un răspuns ușor sau corect; fiecare caz va genera victime. Dar trebuie dat un răspuns, oricât de greu ar fi; iar aceasta se poate face doar înțelegând cum funcționează mașinile autonome și ajustându-ne, prin urmare, așteptările față de ele.

Datorită profesiei, sunt eu însumi un creator de algoritmi care, conform intențiilor mele, vor îmbunătăți viețile oamenilor. Motivul pentru care, anterior, menționam uimirea provocată de anumiți ani este că, în domeniul algoritmilor, lucrurile evoluează într-un ritm amețitor. Inovații care cu doar trei ani în urmă făceau senzație pentru că au întrecut performanța umană, azi sunt considerate învechite și au fost întrecute de câteva ori. Din acest punct de vedere, după cum recunoaște și autoarea, opera de față va deveni, poate, incompletă de îndată ce va ieși din tipografie. Totuși, consider că *Hello, world!* reprezintă cel mai cuprinzător studiu al impactului pe care îl au algoritmii. După cum veți vedea, cele șapte capitole tratează influența algoritmilor în tot atâtea

* Poate mai puțin cel către melodia formației Vanilla, care, amintindu-ne, cumva, de epoca lui „Lasă-mă, papa, la mare!", este un kitsch melodic ce a fost împins forțat în topurile radio din Marea Britanie.

** Vezi capitolul „Automobilele autonome".

domenii, care sunt considerate, în principal, tipic umane, spre exemplu medicina, justiția și chiar arta. Și în nenumărate rânduri sunt prezentate cazuri în care algoritmii surclasează oamenii. Dar, oare, înseamnă asta că ne pot înlocui? M-am oprit deseori din citit pentru a reflecta la acest lucru.

Se spune, în ziua de azi, că inteligența artificială provoacă deja cea de-a patra revoluție industrială*, iar lideri din întreaga lume își fac griji în legătură cu impactul pe care îl va avea asupra societății, oferind ca motiv principal viteza cu care aceasta se desfășoară. Ușurința cu care companiile vechi și noi se pot extinde la scară globală le conferă un avantaj fără precedent, iar legislatorii se chinuie să țină pasul cu tehnologia. În plus, aplicabilitatea largă a acestor algoritmi, care include, după cum am menționat, domenii tipic umane, inspiră motive de îngrijorare cu privire la locurile noastre de muncă. Am asistat de curând la o prezentare a lui Garry Kasparov care, după cum veți afla din primul capitol, are motive foarte întemeiate să poarte pică mașinilor cu performanțe super-umane. Am fost surprins să aflu că, în schimb, el este optimist și susține ideea dezvoltării unor mașini din ce în ce mai performante. „Când au apărut tractoarele, spunea el, vuiau ziarele de știri bucuroase că o astfel de mașină poate înlocui munca asiduă a 30 de oameni. Nimeni nu era ofensat că un tractor era mai puternic decât el. Ba, mai mult, oamenii se înghesuiau să învețe să le conducă. De ce lucrurile ar fi diferite acum?"

Este îndreptățit, oare, Garri să fie atât de optimist în privința algoritmilor? Vă invit să vă formați propria opinie cu ajutorul acestei lucrări, care va sta, cred eu, la temelia înțelegerii puterii mașinilor din secolul al XXI-lea și, totodată, a evaluării failibilității oamenilor.

CIPRIAN I. TOMOIAGĂ

* Forumul Economic Mondial, https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond (link accesat cel mai recent la 19 martie 2019).

Dedicată Mariei Fry.

Îţi mulţumesc pentru că nu ai acceptat
niciodată să ţi se spună „nu"!

NOTĂ DESPRE TITLU

Când aveam șapte ani, tata ne-a adus, mie și surorilor mele, un cadou: un ZX Spectrum, un mic computer de numai opt biți. Pentru prima oară, aveam un calculator numai și numai al nostru! Probabil că era deja perimat de vreo cinci ani în momentul în care a ajuns la noi, însă, chiar dacă era doar un computer la mâna a doua, am simțit imediat că mașina* aceea simpatică era un lucru cu totul și cu totul minunat. Spectrum era, oarecum, echivalentul unui Commodore 64 (cu toate că numai copiii fițoși din cartier aveau așa ceva), deși eu am considerat întotdeauna că Spectrum era o drăcovenie mult mai frumoasă. Puteai ține foarte comod în mâini carcasa netedă din plastic negru, iar tastele din cauciuc gri și dunga în culorile curcubeului așezată în diagonală peste unul dintre colțuri îi dădeau un aer destul de prietenos.

În ceea ce mă privește, sosirea acelui ZX Spectrum a însemnat începutul unei veri memorabile, pe care am petrecut-o în mansardă, împreună cu sora mea mai mare, programând una pentru cealaltă jocuri de-a spânzurătoarea sau desenând forme geometrice simple cu ajutorul unui cod.

* *Machine* (engl.), „(aici) computer". Aici și peste tot în continuare am optat să traducem acest termen în conformitate cu definiția din Traian Anghel, *Dicționar de informatică* (Editura Corint, București, 2017): „termen echivalent cu calculator, gazdă sau sistem" (pentru o definiție mai detaliată, vezi glosarul, p. 318) (n. red.).

Toate aceste chestiuni „avansate" au venit, însă, mai târziu. Pentru început, trebuia să stăpânim bine noțiunile de bază.

Privind în trecut, nu-mi amintesc cu precizie momentul în care am scris primul meu program de computer, dar sunt destul de sigură că știu despre ce era vorba. E posibil să fi fost același program simplu pe care am ajuns să-l predau tuturor studenților mei de la University College London (UCL), același pe care-l veți găsi pe prima pagină a oricărui manual de introducere în informatică*. Pentru că s-a împământenit o tradiție printre toți cei care învață să programeze, una care e, aproape, un fel de ritual de inițiere. Prima însărcinare a unui novice în ale computerelor este să programeze calculatorul să afișeze pe ecran o expresie bine-cunoscută:

HELLO, WORLD!

Este o tradiție care datează încă din anii 1970, când Brian Kernighan a inclus-o ca lucrare practică în manualul său de programare extrem de cunoscut.[1] Cartea – de unde provine expresia – a marcat un moment important în istoria calculatoarelor. Microprocesorul tocmai își făcuse apariția în scenă, anunțând transformarea computerelor din ceea ce fuseseră în trecut – niște mașini grozave pentru specialiști, alimentate cu cartele și benzi perforate – în ceva mai asemănător cu PC-urile pe care le cunoaștem, cu ecran, tastatură și cursor care clipește. *Hello, world!* și-a făcut apariția în momentul în care a devenit o posibilitate să comunici cu computerul.

Ani buni mai târziu, Brian Kernighan i-a relatat unui redactor al revistei *Forbes* ce l-a inspirat să aleagă această expresie. Văzuse un desen animat cu un ou și cu un puișor abia ieșit din găoace, care, în clipa în care se născuse, ciripise cuvintele: *Hello, world!* și imaginea acesta îi rămăsese în minte.

* Pentru definițiile unor termeni de specialitate, aici și în continuare, vezi glosarul, p. 315 (n. red.).

Nu se ştie prea bine cine era puişorul în tot acest scenariu: tinerii care îşi anunţau, triumfători, pătrunderea curajoasă în universul programării computerelor? Sau PC-ul însuşi, trezit din moţăiala sa plictisită printre foi de calcul şi documente-text, pregătit să-şi conecteze mintea la lumea reală şi să îndeplinească cererile noilor săi stăpâni? Poate că amândouă... Cert e, însă, că aceasta rămâne, cu siguranţă, o expresie care îi uneşte pe toţi programatorii şi îi pune în legătură cu fiecare computer care a fost programat vreodată.

Şi trebuie să mărturisesc că îmi mai place ceva la expresia aceasta – ceva ce nu a fost niciodată mai relevant sau mai important decât este în prezent. Căci, în timp ce algoritmii informatici ne controlează şi ne decid viitorul într-o măsură din ce în ce mai mare, *Hello, world!* ne reaminteşte de un moment în care între om şi computer se purta un dialog. Ne reaminteşte de clipa în care graniţa care îl desparte pe cel care controlează de cel controlat devine, practic, insesizabilă. Şi marchează, totodată, începutul unui parteneriat – un parcurs comun de explorare a posibilităţilor, în care unul nu poate exista fără de celălalt.

Iar, în era maşinilor, acesta este un sentiment de care merită să ţinem seama.

INTRODUCERE

ORICINE A VIZITAT VREODATĂ parcul național Jones Beach din Long Island, statul american New York, probabil că a trecut cu automobilul, în drum spre ocean, pe sub câteva poduri. Acestea au fost construite, inițial, în folosul celor care intrau sau ieșeau de pe autostradă și au o caracteristică neobișnuită. Arcuindu-se ușor pe deasupra autostrăzii, ele se află la o înălțime extrem de mică de aceasta, lăsând uneori o distanță de numai 2,74 metri până la asfalt.

Designul ciudat are o explicație, însă. În anii 1920, Robert Moses, un urbanist bogat și puternic din New York, a dorit ca recent terminatul și foarte premiatul parc-rezervație de la Jones Beach să le fie accesibil numai și numai americanilor albi înstăriți. Și, cum știa prea bine că toți clienții săi preferați aveau să se deplaseze până la plajă în autoturismele lor personale, în vreme ce locuitorii de culoare din cartierele sărace n-aveau să poată ajunge acolo decât cu autobuzul, el a încercat, în mod intenționat, să le limiteze celor din urmă posibilitatea de a merge acolo, construind sute de poduri joase de-a lungul autostrăzii. Atât de joase, încât autobuzelor înalte de 3,5 metri să le fie imposibil să treacă pe sub ele.[1]

Podurile acestea de-a dreptul rasiste nu sunt singurele obiecte care, deși neînsuflețite, au reușit să exercite un control tăcut, clandestin, asupra ființelor umane. Istoria este pângărită de nenumărate obiecte și invenții care posedau o

putere superioară scopului lor declarat.[2] Uneori, aceasta fusese integrată, în mod deliberat și răuvoitor, în însuși conceptul lor, de la bun început, dar alteori nu a fost decât rezultatul unor omisiuni nechibzuite: gândiți-vă un pic numai la inexistența, în unele zone urbane, a rampelor de acces pentru persoanele țintuite în scaune cu rotile! Iar câteodată, a fost, pur și simplu, o consecință neintenționată, așa cum a fost cazul războaielor de țesut din secolul al XIX-lea, care fuseseră fabricate pentru a ușura obținerea unor materiale textile complexe, însă, în cele din urmă, au ajuns să aibă un impact uriaș asupra salariilor, șomajului și condițiilor de lucru, încât au devenit, în mod indiscutabil, mai tiranice decât orice capitalist din epoca victoriană.

Invențiile moderne nu sunt, nici ele, diferite. Întrebați-i numai pe locuitorii din orășelul Scunthorpe* din nordul Angliei ce probleme au avut când li s-a interzis să-și deschidă conturi pe AOL după ce uriașul furnizor de internet a pus la punct un nou filtru, care excludea vulgaritățile și care a avut obiecții față de însăși denumirea localității lor.[3] Sau pe Chukwuemeka Afigbo, nigerianul care a descoperit că un dozator automat pentru săpun lichid elibera soluția, în mod impecabil, atunci când prietenul său alb își punea mâinile sub aparat, însă refuza, cu obstinație, să facă același lucru în cazul său, nerecunoscând nuanța mai închisă a pielii sale.[4] Sau pe Mark Zuckerberg, care, atunci când concepea codul pentru Facebook în camera lui de student de la Harvard, în 2004, nu și-ar fi imaginat niciodată că propria creație avea să fie acuzată că a înlesnit manipularea voturilor în alegeri desfășurate pe întregul mapamond.[5]

* În numele orășelului se distinge cuvântul *cunt* (engl.), „(trivial, peiorativ și ofensator) vagin". În limbajul curent a fost introdusă expresia *Scunthorpe problem*, care desemnează o dificultate ce apare atunci când filtrele de internet blochează nume legitime pentru că acestea conțin anumite combinații de litere, care, scoase din context, pot fi considerate obscene (n. tr.).

În spatele fiecăreia dintre aceste invenții stau algoritmii. Fragmente invizibile de cod care alcătuiesc mecanismele și rotițele mașinilor epocii moderne, algoritmii au oferit lumii totul, de la fluxurile de știri ale mass-media și până la motoarele de căutare, de la navigarea prin satelit și până la programele de recomandare a muzicii, făcând tot atât de mult parte din infrastructura noastră modernă pe cât au făcut, cândva, și podurile, clădirile sau fabricile. Există algoritmi în spitale, în sălile de judecată și în autovehiculele pe care le conducem. Tot ei sunt folosiți de forțele de poliție, de către lucrătorii din supermarketuri, dar și de către angajații studiourilor cinematografice. Algoritmii au învățat ce ne place și ce ne displace; ei ne spun ce să urmărim, ce să citim și cu cine să avem o întâlnire romantică. Iar, în tot acest timp, au ajuns să dețină puterea ascunsă de a schimba, încet-încet și în mod subtil, regulile în privința a ceea ce înseamnă să fii om.

În această carte vom descoperi marea varietate de algoritmi pe care ne bazăm din ce în ce mai mult, poate chiar fără să ne dăm seama. Ne vom îndrepta atenția asupra a ceea ce pretind ei că fac, le vom analiza puterea nedeclarată și vom încerca să ne confruntăm cu întrebările rămase fără răspuns pe care le ridică. Vom face cunoștință cu algoritmii folosiți de către polițiști pentru a decide cine să fie arestat, pe aceia care ne fac să alegem dacă să le protejăm pe victimele infracționalității sau să credem în inocența celor care sunt puși sub acuzare. Vom afla care sunt algoritmii pe care îi folosesc judecătorii pentru a stabili sentințele date infractorilor care au fost găsiți vinovați, adică acelea care ne somează, practic, să decidem cum ar trebui să fie sistemul nostru judiciar. Vom descoperi că există algoritmi pe care medicii îi folosesc pentru a-și reexamina sau chiar pentru a-și infirma propriile diagnostice, algoritmi în interiorul autoturismelor fără șofer care tot insistă să ne reevaluăm propria moralitate, dar și unii care cântăresc modul în care ne exprimăm emoțiile, precum și alții care dețin puterea de a ne submina democrațiile.

Nu susțin, nici pe departe, că algoritmii în sine ar fi nocivi. Așa cum veți vedea în paginile următoare, există nenumărate motive pentru care merită să rămânem pozitivi și să privim cu optimism spre ceea ce ne așteaptă în viitor. Tot ce contează este modul în care sunt utilizați algoritmii. GPS-ul a fost inventat pentru lansarea rachetelor nucleare, iar acum ajută la livrarea pizzei, câtă vreme s-a descoperit că repetarea la nesfârșit a unei melodii pop poate fi folosită ca metodă de tortură. Totodată, să știți că, oricât de frumos va fi fost împletită o ghirlandă de flori, dac-aș dori neapărat acest lucru, chiar aș putea folosi una pentru a strangula pe cineva! Să îți formezi o opinie în privința unui algoritm înseamnă să înțelegi relația care există între om și mașină. Fiecare algoritm este, în mod obligatoriu, legat de persoanele care l-au conceput și îl utilizează.

Ceea ce înseamnă că, în esență, aceasta este o carte despre oameni. Despre felul în care suntem, despre direcția în care ne îndreptăm, despre ceea ce este important pentru noi și despre felul în care toate acestea sunt transformate prin intermediul tehnologiei. Este o carte despre relația pe care o avem cu algoritmii care există deja, cu aceia care lucrează împreună cu noi, ne potențează abilitățile, ne corectează greșelile, ne rezolvă problemele, deși ne mai creează și altele noi pe parcurs.

Aici ne vom pune întrebări precum: aduc sau nu algoritmii beneficii certe societății? Când ar trebui să ai mai multă încredere într-o mașină decât în propria putere de judecată și când ar trebui să reziști tentației de a lăsa computerele să comande? Vom descompune algoritmii și le vom descoperi limitările, iar, în cele din urmă, ne vom cerceta cu atenție pe noi înșine și le vom descoperi pe ale noastre. În această carte vom încerca să separăm ce e rău de ce e bine și să hotărâm în ce fel de lume dorim să trăim.

Pentru că viitorul nu e ceva ce ni se întâmplă, pur și simplu. Noi suntem aceia care avem puterea de a-l crea.

Capitolul 1

PUTEREA

Garry Kasparov ştia exact cum să-şi intimideze adversarii. La 34 de ani, era cel mai mare jucător de şah care existase vreodată, având o reputaţie suficient de redutabilă, încât să îl facă pe orice rival al său să fie cuprins de neliniște. Chiar şi aşa, apela la un şiretlic enervant pe care adversarii săi ajunseseră să-l urască. În timp ce erau aşezaţi, asudând să câştige ceea ce era, probabil, cea mai grea partidă din viaţa lor, rusul îşi lua, nonşalant, ceasul de mână de unde fusese pus, lângă tabla de şah, şi şi-l fixa din nou la încheietură. Era semnalul pe care îl recunoştea toată lumea, acela care însemna că Garry Kasparov se plictisise să se joace cu adversarul său. Ridicarea ceasului semnala că era timpul ca rivalul său să abandoneze partida. Acesta putea să refuze, însă Kasparov avea să-l învingă oricum în curând, iar acest lucru era inevitabil[1].

Dar atunci când Deep Blue al lui IBM* l-a înfruntat pe Kasparov, în meciul celebru din mai 1997, maşina era imună la asemenea tactici. Rezultatul partidei este bine-cunoscut, însă povestea modului în care Deep Blue şi-a asigurat victoria este

* În mai 1997, Gari (sau Garry) Kasparov a jucat la New York un meci de şase partide cu super-computerul IBM, Deep Blue. Evenimentul a fost urmărit de milioane de oameni, în timp real, pe site-ul IBM. Dintre cele şase partide desfăşurate pe parcursul a şapte zile, Kasparov nu a reuşit să învingă calculatorul decât în prima partidă (n. tr.).

mai puțin știută. Acea triumfare simbolică a computerului asupra ființei umane, care a marcat, în multe privințe, începutul erei algoritmilor, nu s-a datorat doar simplei puteri de calcul superioare. Pentru a-l învinge pe Kasparov, Deep Blue fusese nevoit să-l înțeleagă nu doar ca pe un procesor extrem de eficient pentru calcularea unor mutări de șah senzaționale, ci și ca pe o ființă umană.

Pentru început, inginerii de la IBM au luat geniala decizie de a-l proiecta pe Deep Blue astfel încât să pară mult mai nesigur decât era în realitate. Pe durata celebrului meci format din șase partide, mașina s-a abținut, din când în când, să anunțe ce mutarea dorea să efectueze după ce își terminase deja calculele, uneori chiar și pentru câteva minute bune. Din perspectiva lui Kasparov, întârzierile păreau să lase de înțeles că mașina se căznea, chinuindu-se să efectueze din ce în ce mai multe calcule. Asta părea să confirme ceea ce Kasparov credea că știa: și anume, că reușise să aducă jocul într-o situație în care numărul posibilităților era atât de năucitor de mare, încât Deep Blue nu putea să ia o hotărâre judicioasă[2]. În realitate, acesta stătea în repaus, știind exact ce piesă să mute și lăsând, pur și simplu, timpul să se scurgă. Era un șiretlic plin de viclenie, dar a funcționat. Chiar din prima partidă a meciului, Kasparov a început să se lase din ce în ce mai distras, tot încercând să aprecieze cât de priceput la șah putea fi Deep Blue[3].

Deși Kasparov a câștigat prima partidă, de la a doua, Deep Blue a reușit să i se vâre pe sub piele, deslușindu-i modul de a gândi. Kasparov a încercat să-l ademenească într-o capcană, ispitindu-l să captureze câteva piese, în timp ce el însuși se pregătea – cu câteva mișcări în avans – să-și elibereze regina și să lanseze un atac[4]. Toți șahiștii care urmăreau partida, ca și Kasparov însuși, se așteptau ca Deep Blue să muște momeala și să cadă în capcană. Numai că, nu se știe cum, computerul a mirosit el ceva. Spre uimirea lui Kasparov, Deep Blue și-a dat seama ce punea la cale marele maestru și i-a

blocat acestuia regina, distrugându-i, astfel, orice şansă de a obţine o victorie umană[5].

Kasparov a fost, în mod vizibil, oripilat. Faptul că se înşelase în legătură cu ceea ce era computerul capabil să facă l-a terminat, pur şi simplu. Câteva zile mai târziu, într-un interviu luat după meci, el a apreciat că Deep Blue jucase „dintr-odată, pentru o clipă, ca un zeu[6]". Mulţi ani mai târziu, reflectând la cum se simţise atunci, Kasparov a scris că „făcuse greşeala să creadă că mutările pe care era surprinzător ca un calculator să le facă reprezentau, totodată, şi nişte mutări grozave, imbatabile, în mod obiectiv".[7] În orice caz, geniul algoritmului triumfase. Felul în care reuşise să înţeleagă mintea omenească şi propensiunea acesteia spre greşeală îl ajutase să îl atace şi să îl învingă pe geniul cel mult prea uman.

Demoralizat, Kasparov a preferat să abandoneze a doua partidă, în loc să lupte pentru a obţine o remiză. Din acel moment, încrederea lui în sine a început să se destrame.

A treia, a patra şi a cincea partidă s-au soldat cu remize. Când s-a ajuns la a şasea, Kasparov era deja terminat. Scorul final al meciului: Deep Blue, 3½, Kasparov, 2½.

A fost o înfrângere ciudată. Kasparov era mai mult decât capabil să iasă din toate poziţiile dificile în care îi ajunseseră piesele de pe tabla de şah, însă el subestimase capacitatea algoritmului, iar apoi se lăsase intimidat de acesta. Reflectând la acel meci, el avea să scrie în 2017:

> Am fost atât de impresionat de felul în care juca Deep Blue! Am devenit atât de preocupat de ceea ce era capabil să facă, încât nu am mai izbutit să-mi dau seama că toate problemele mele se datorau mai mult faptului că eu jucam prost, decât faptului că el juca prea bine.[8]

Aşa cum vom vedea în repetate rânduri în această carte, aşteptările, speranţele noastre, sunt importante. Povestea înfrângerii unui mare maestru de şah de către Deep Blue arată că puterea unui algoritm nu se limitează neapărat la ceea ce

conțin liniile sale de cod. Înțelegerea propriilor defecte și slăbiciuni – precum și pe acelea ale computerului – reprezintă cheia pentru deținerea controlului.

Dar dacă un geniu precum Kasparov nu a înțeles acest lucru, ce speranțe mai avem noi, toți ceilalți? În aceste pagini vom vedea cum algoritmii s-au strecurat, practic, în absolut toate aspectele vieții moderne – de la serviciile medicale și lupta împotriva infracționalității, până la transport și politică. Am reușit, cumva, în decursul timpului, și să-i ignorăm, și să ne lăsăm intimidați de ei, dar și să ne minunăm de capacitățile lor, toate în același timp. Rezultatul este că nu prea avem habar cât de multă putere le conferim și nici dacă nu cumva am îngăduit deja ca lucrurile să meargă prea departe.

ÎNAPOI LA ELEMENTELE FUNDAMENTALE

Înainte de a ajunge să discutăm despre toate acestea însă, ar fi bine, poate, să ne oprim puțin și să ne întrebăm ce înseamnă, mai exact, un *algoritm*. Este o noțiune care, deși frecvent utilizată, nu reușește, cel mai adesea, să transmită prea multe. Asta se întâmplă, în parte, deoarece termenul în sine este destul de neclar. Oficial, el este definit după cum urmează[9]:

> *algoritm* (s.m.), o procedură pas cu pas de soluționare a unei probleme sau de îndeplinire a unei sarcini, în principal de către un computer.

Asta e tot! Un algoritm reprezintă, pur și simplu, o serie de instrucțiuni logice care arată, de la început și până la sfârșit, cum să se îndeplinească o anumită sarcină. Conform acestei definiții generale, o rețetă pentru prăjituri poate fi considerată un algoritm. La fel și o succesiune de îndrumări pe care i le-ai oferi unui străin care s-a rătăcit. Manualele de la IKEA, videoclipurile despre cum să repari diverse chestii de pe YouTube, până și cărțile de tip *self-help* – teoretic, orice listă de instrucțiuni

autonome cu ajutorul cărora se poate atinge un obiectiv specific şi bine definit ar putea fi considerată un algoritm.

Numai că termenul nu este utilizat chiar în acest sens. De regulă, când spunem *algoritmi* ne referim la ceva special: încă mai reprezintă, în esenţă, nişte liste de instrucţiuni care trebuie urmate pas cu pas, dar aceşti algoritmi sunt, aproape întotdeauna, nişte obiecte matematice. Ei preiau secvenţe de operaţiuni matematice şi (folosind ecuaţii, calcule aritmetice sau algebrice, ori chiar analiza matematică, logica şi probabilităţile) le traduc în cod-program. Algoritmii sunt „alimentaţi" cu informaţii din lumea reală, li se atribuie un obiectiv şi sunt puşi la lucru, să calculeze şi să tot calculeze până îşi ating scopul. Algoritmii sunt aceia care fac din informatică o adevărată ştiinţă, iar, în decursul timpului, au facilitat multe dintre cele mai extraordinare creaţii moderne realizate de diverse maşini.

Există un număr aproape incalculabil de algoritmi diferiţi. Fiecare cu propriile obiective, propriile caracteristici, ingenioasele sale ciudăţenii şi deficienţe, aşa că nu s-a ajuns încă la un consens cu privire la cel mai bun mod de a-i clasifica. Dar, în mod general, poate fi util să ne gândim la sarcinile din lumea reală pe care le îndeplinesc algoritmii, grupându-le în patru mari categorii[10]:

Stabilirea priorităţilor: conceperea unei liste ordonate

Google Search anticipează ce site cauţi, deoarece ierarhizează, stabileşte prioritatea unui site pe care s-ar putea să doreşti să-l accesezi în funcţie de frecvenţa cu care l-ai utilizat în trecut, comparând rezultatele şi clasificându-le. În mod similar procedează Netflix când îţi sugerează filme pe care s-ar putea să-ţi placă să le urmăreşti sau TomTom* când îţi alege ruta cea mai rapidă. Toate folosesc un proces matematic pentru a ordona marea varietate de alegeri posibile. Deep Blue a fost, în principal, tot un algoritm de stabilire

* Companie fondată în 1991, cu sediul central în Amsterdam, Olanda, specializată în sisteme de navigare auto şi hărţi (n. tr.).

de priorități, unul care trecea în revistă toate mutările posibile de pe tabla de șah și calcula care dintre ele ar putea oferi șansa cea mai bună de a învinge.

Clasificarea: alegerea unei categorii

Când m-am apropiat de vârsta de 30 de ani, am fost bombardată pe Facebook de reclame la inele de logodnă, împodobite cu diamante. Iar, de îndată ce m-am căsătorit, reclamele pentru testele de sarcină mă urmăreau peste tot pe internet. Pentru aceste mici sâcâieli trebuie să le mulțumesc algoritmilor de clasificare. Aceștia, care sunt atât de îndrăgiți de cei care lucrează în agențiile de publicitate, rulează în spatele scenei și te clasifică drept o persoană interesată de lucrurile respective pe baza caracteristicilor personale. (Și, cu toate că s-ar putea să aibă dreptate, tot este supărător când reclamele la teste de fertilitate îți tot răsar pe laptop în mijlocul unei ședințe!)

Există algoritmi care pot să clasifice și să îndepărteze automat conținutul inadecvat de pe YouTube, alții care pot eticheta pozele de vacanță în locul tău și algoritmi care îți pot scana scrisul de mână și identifica orice semn de pe o pagină drept o literă din alfabet.

Asocierea: găsirea legăturilor

Asocierea se referă la găsirea și evidențierea relațiilor dintre lucruri. Algoritmii pentru site-urile de anunțuri matrimoniale, precum OKCupid, se bazează pe asociere, ei căutând legături între membri și sugerând potriviri în funcție de rezultatele găsite. Motorul de recomandări al companiei Amazon folosește o idee similară, făcând asocieri între ceea ce te interesează și ceea ce i-a interesat pe foștii clienți. Aceasta a dus la uimitoarea sugestie de cumpărături care i s-a făcut utilizatorului *Kerbobotat* de pe site-ul Reddit, după ce cumpărase o bâtă de baseball pe Amazon: „Poate vă interesează să cumpărați această cagulă?”[11]

Filtrarea: izolarea a ceea ce este important

Algoritmii trebuie adesea să îndepărteze o parte din informații pentru a se concentra asupra a ceea ce este important, să separe, adică, semnalul de zgomot. Uneori fac asta *ad litteram*: astfel, algoritmii de recunoaștere a vorbirii, cum sunt cei care rulează în interiorul Siri, Alexa și Cortana, trebuie, mai întâi, să-ți filtreze vocea, pentru a o deosebi de zgomotele din fundal înainte de a începe să lucreze la descifrarea spuselor tale. Uneori o fac în sens figurat: Facebook și Twitter filtrează, în funcție de activitatea ta anterioară, postările care au legătură cu ceea ce te interesează, pentru a concepe un flux de știri (*feed*) personalizat, care ți se adresează numai și numai ție.

Majoritatea algoritmilor sunt astfel concepuți, încât să execute o combinație a sarcinilor de mai sus. Spre exemplu, UberPool potrivește, în mod automat, pasagerii potențiali cu alții, care se deplasează în aceeași direcție. Odată ce i se dau punctul de plecare și destinația, aplicația trebuie să selecteze posibilele rute care te-ar duce acasă, să te asocieze cu alți utilizatori care se îndreaptă în aceeași direcție și să te aloce unui grup pe care l-a ales – toate acestea, în timp ce selectează, cu prioritate, rutele care vor obliga șoferul să execute cele mai puține schimbări de direcție, pentru a face cursa cât mai eficientă cu putință.[12]

Prin urmare, asta pot face algoritmii. Acum, să vedem cum reușesc să ducă la îndeplinire toate acestea. Din nou, deși posibilitățile sunt, practic, infinite, există un mod de a sintetiza lucrurile. Putem considera că abordările utilizate de algoritmi se încadrează, în linii mari, în două paradigme principale pe care le vom întâlni în această carte.

Algoritmi care se bazează pe reguli

Primul tip de algoritmi se bazează pe reguli. Instrucțiunile lor sunt scrise de un operator uman, sunt directe și lipsite de ambiguitate. Ne putem imagina că algoritmii aceștia urmează o schemă logică precum rețeta unei prăjituri. Primul pas:

fă asta. Al doilea pas: dacă faci asta, atunci rezultă cealaltă. Ceea ce nu înseamnă deloc că aceşti algoritmi ar fi simpli – există suficient spaţiu pentru construirea unor programe puternice în cadrul amintitei paradigme.

Algoritmi de învăţare automată

Al doilea tip de algoritmi s-a inspirat din modul în care învaţă fiinţele vii. Pentru a oferi o analogie, să ne gândim cum am putea să-l învăţăm pe un câine să ne dea laba. Nu este nevoie să întocmim o listă exactă de instrucţiuni pe care să i le comunicăm. Dresorul nu trebuie decât să îşi stabilească bine în minte un obiectiv (ce doreşte să facă acel câine), precum şi o modalitate oarecare de a-l răsplăti atunci când duce la îndeplinire în mod corect sarcina respectivă. Totul se rezumă, cât se poate de simplu, la consolidarea comportamentului dezirabil şi la ignorarea celui nedorit, şi, desigur, la exersare, astfel încât câinele să ajungă să-şi dea singur seama ce are de făcut. Când vorbim despre algoritmi, echivalentul celor de mai sus este cunoscut sub denumirea de „algoritm de învăţare automată", care se înscrie în categoria mai largă a „inteligenţei artificiale", sau IA*. Îi furnizezi maşinii informaţii, un obiectiv şi feed-back atunci când se află pe calea cea bună – şi o laşi să descopere singură modalitatea cea mai bună de a-şi atinge scopul.

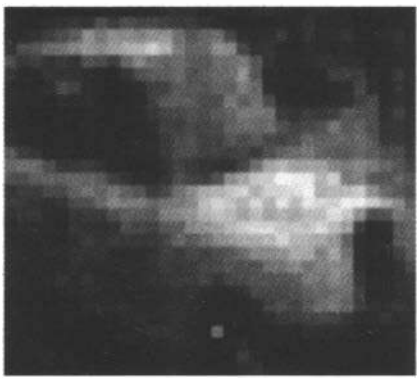
Ambele tipuri de algoritmi au părţi pozitive şi negative. Din moment ce algoritmii care se bazează pe reguli au instrucţiuni scrise de fiinţe umane, ei sunt uşor de înţeles. Teoretic, oricine îi poate descompune şi urmări raţionamentul a ceea ce se petrece înăuntrul lor.[13] Însă această binecuvântare este, pentru ei, şi un blestem. Algoritmii care se bazează pe reguli vor funcţiona numai pentru rezolvarea problemelor pentru care operatorii umani ştiu cum să scrie instrucţiunile.

Algoritmii de învăţare automată, în schimb, s-au dovedit, recent, a fi remarcabil de buni la rezolvarea unor probleme pentru care scrierea unei liste de instrucţiuni nu este posibilă.

* *AI, artificial intelligence* (engl.) (n. tr.).

Ei pot recunoaşte obiecte în diverse imagini, ne pot înţelege cuvintele pe măsură ce le rostim şi pot să traducă dintr-o limbă în alta – ceva ce a fost întotdeauna dificil pentru algoritmii care se bazează pe reguli. Dezavantajul este că, dacă vei lăsa o maşină să descopere singură soluţia, procesul pe care îl va urma pentru a-şi îndeplini obiectivul nu va avea sens pentru un observator uman. Căci părţile sale tainice pot fi o enigmă chiar şi pentru cei mai pricepuţi programatori ai zilelor noastre.

Să luăm, de exemplu, sarcina recunoaşterii unei imagini. Un grup de cercetători japonezi a demonstrat de curând cât de bizar i se poate părea unui om modul în care un algoritm „vede" lumea. Poate că aţi mai întâlnit iluzia optică în care nu puteţi spune cu certitudine dacă vă uitaţi la imaginea unei vaze sau la aceea a două chipuri umane (în caz contrar, aveţi un exemplu în notele de la sfârşitul cărţii[14]). Iată care ar fi echivalentul acesteia pentru un computer: echipa de cercetători a arătat că simpla modificare a unui singur pixel la roata din faţă a imaginii următoare fusese de ajuns pentru a face ca un algoritm de învăţare automată să se răzgândească şi să considere că fotografia nu mai reprezenta un automobil, ci un câine.[15]



Pentru unii, ideea ca un algoritm să lucreze fără instrucțiuni explicite pare să reprezinte cea mai sigură cale spre dezastru. Cum putem controla ceva ce nu înțelegem? Ce se va întâmpla în cazul în care capacitățile mașinilor conștiente de sine și super-inteligente le vor depăși pe acelea ale creatorilor lor? Cum am putea garanta că o inteligență artificială (IA) pe care nu o înțelegem și pe care nu o putem controla nu complotează deja împotriva noastră?

Toate acestea sunt întrebări ipotetice interesante și nu ducem lipsă de cărți despre amenințarea iminentă a unei apocalipse cauzate de IA. Îmi cer scuze dacă la așa ceva vă așteptați, însă cartea de față nu este una dintre ele. Cu toate că IA a făcut progrese rapide în ultima vreme, aceasta rămâne „inteligentă" numai în sensul cel mai restrâns al noțiunii. Ar fi, poate, mai util să ne gândim că am asistat la o revoluție în ceea ce privește statistica informatizată, mai degrabă decât la o revoluție în crearea inteligenței artificiale. Știu că sună mult mai puțin atrăgător (dacă nu cumva *chiar* sunteți pasionați de statistică!), însă este o descriere mult mai corectă a modului în care stau lucrurile în prezent.

Pentru moment, să ne facem griji din cauza unei IA malefice pare, oarecum, ca și cum ne-am face griji în legătură cu... suprapopularea planetei Marte!* Poate că într-o bună zi vom ajunge și în punctul în care inteligența artificială o va depăși pe cea umană, însă, pentru moment, nu suntem nici pe departe acolo. La drept vorbind, suntem încă departe de crearea fie și a unei inteligențe de nivelul celei pe care o are un arici. Până

* Am parafrazat aici un comentariu făcut de Andrew Ng, specialistul în calculatoare și pionierul *deep learning*, într-o conferință pe care a susținut-o în 2015. Vezi Tech Events, GPU Technology Conference 2015, ziua a III-a: „What's Next in Deep Learning", discurs postat pe YouTube (https://www.youtube.com/watch?v=qP9TOX8T-kI) la 20 noiembrie 2015 (link accesat cel mai recent la 4 martie 2019) (n. a.).

acum, nimeni nu a reușit să depășească nivelul inteligenței unui vierme.*

În plus, toată agitația stârnită în legătură cu IA reprezintă o abatere a atenției de la preocupări mult mai urgente și – cred eu – de la subiecte mult mai interesante. Haideți, așadar, să dăm uitării, pentru moment, omnipotenta inteligență artificială și să ne întoarcem, din viitorul foarte îndepărtat, la timpul prezent, pentru că deja există algoritmi cărora li s-a dat frâu liber să acționeze și să ia, în mod autonom, decizii. Unii care stabilesc perioadele de încarcerare ale infractorilor condamnați, tratamentele pe care trebuie să le urmeze bolnavii de cancer și ce trebuie făcut în cazul unui accident rutier. Pentru că algoritmii iau deja, în numele nostru, hotărâri care ne schimbă viețile.

Întrebarea este: dacă le oferim toată această putere, oare, merită ei toată încrederea noastră?

ÎNCREDERE OARBĂ

Duminică, 22 martie 2009, nu a fost o zi bună pentru Robert Jones. Tocmai își vizitase niște prieteni și se întorcea cu mașina prin frumosul oraș Todmorden din West Yorkshire, când a observat că i s-a aprins becul indicatorului de benzină al BMW-ului. Mai avea doar 11 kilometri până să găsească o stație de benzină înainte să i se termine combustibilul, ceea ce era chiar la limită. Din fericire, GPS-ul părea să-i fi găsit o

* Simularea creierului unui vierme constituie chiar obiectivul proiectului științific internațional OpenWorm. Cercetătorii speră să reproducă, pe cale artificială, rețeaua formată din cei 302 neuroni care se găsesc în creierul unui vierme *Caenohabditis elegans*. Pentru a pune lucrurile în perspectivă, trebuie menționat că noi, oamenii, avem în jur de 100 000 000 000 de neuroni. Pentru mai multe informații, vezi site-ul proiectului OpenWorm: http://openworm.org/ (link accesat cel mai recent la 4 martie 2019) (n. a.).

scurtătură și l-a trimis pe un drumeag îngust și sinuos care urca pe coasta unui deal.

Robert a urmat indicațiile mașinii, numai că, pe măsură ce conducea, drumul devenea tot mai abrupt și mai îngust, până când, după vreo trei kilometri, s-a transformat într-un adevărat drum de țară pe care abia-abia dac-ar fi încăput niște cai, nicidecum automobile! Dar Robert nu a fost impresionat. Conducea câte 8 000 de kilometri pe săptămână ca să-și câștige pâinea și știa cum să se descurce când era la volan. În plus, s-a gândit el, nu avea „niciun motiv să nu aibă încredere în sistemul său de navigare prin satelit TomTom."[16]

Doar câteva clipe mai târziu, oricine s-ar fi întâmplat să privească în sus din vale ar fi văzut botul BMW-ului lui Robert ivindu-se peste marginea stâncoasă a prăpastiei, fiind salvat de căderea în gol de vreo 30 de metri numai de gardul șubred din lemn de care tocmai se izbise.

A fost nevoie de un tractor și trei ATV-uri ca să-i recupereze mașina din locul în care o abandonase. Ceva mai târziu în același an, când s-a prezentat în instanță sub acuzația de conducere imprudentă, Robert a recunoscut că nu se gândise să ignore indicațiile sistemului său de navigare. „Acesta insista că drumeagul acela era o șosea", declarase într-un ziar după incident. „Așa că am avut încredere în el. Nu te aștepți să fii ghidat drept într-o prăpastie!"[17]

Nu, Robert! Nu cred că se așteaptă cineva la una ca asta!

Istorioara aceasta are, cumva, și o morală. Cu toate că s-a simțit, probabil, puțin ridicol, pentru că nu a ținut cont de ceva ce se afla chiar în fața ochilor săi (cum ar fi, bunăoară, o prăpastie văzută prin parbriz!) și pentru că i-a atribuit unui algoritm o inteligență mai mare decât ar fi meritat, Robert Jones nu era, nici pe departe, singurul om aflat în această situație. În fond, Garry Kasparov însuși căzuse în aceeași capcană cu vreo 12 ani în urmă! Și, în moduri mai puțin

spectaculoase, însă la fel de importante, facem aproape toți această greșeală și, probabil, nici măcar nu ne dăm seama.

În 2015, oamenii de știință au început să studieze în ce fel motoarele de căutare precum Google au puterea de a ne modifica perspectiva asupra lumii.[18] Voiau să afle dacă impunem niște limite sănătoase încrederii pe care o acordăm rezultatelor lor sau dacă nu cumva suntem gata să le urmăm, bucuroși, chiar dacă ne-ar ghida, metaforic vorbind, peste marginea unei prăpastii.

Experimentul s-a concentrat asupra alegerilor care urmau să aibă loc în India. Cercetătorii, sub conducerea psihologului Robert Epstein, au selectat 2 150 de alegători indecișI din întreaga țară și le-au oferit acces la un motor de căutare special conceput în acest scop, numit „Kadoodle", pentru a-i ajuta să afle mai multe informații despre candidați înainte de a hotărî pe care să-l voteze.

Kadoodle era „măsluit". Fără ca participanții la studiu să aibă habar, ei fuseseră împărțiți în grupuri, fiecărui grup fiindu-i arătată o versiune ușor diferită a rezultatelor obținute de motorul de căutare, care îl favoriza pe unul sau pe altul dintre candidați. Atunci când membrii unui grup vizitau site-ul, toate link-urile din partea de sus a paginii favorizau un anumit candidat, ceea ce însemna că ei trebuiau să deruleze pagina în jos, link după link, înainte de a găsi fie și o singură trimitere către un site care să îi fie favorabilă oricărui alt candidat. Grupuri diferite au fost încurajate să prefere diverși candidați.

Nu este deloc surprinzător că participanții la experiment și-au petrecut majoritatea timpului citind site-urile semnalate în partea de sus a primei pagini, pentru că, așa cum spune gluma aceea răsuflată, cel mai bun loc unde să ascunzi un cadavru este pe pagina a doua a rezultatelor unei căutări pe Google. Aproape niciunul dintre participanți nu a dat vreo atenție link-urilor care apăreau mult în josul listei. Cu toate acestea, măsura în care ordonarea influența opiniile

voluntarilor l-a şocat până şi pe Epstein. După numai câteva minute de privit rezultatele măsluite ale motorului de căutare, când erau întrebaţi cu cine aveau să voteze, participanţii erau mai dispuşi să aleagă candidatul pe care-l favorizase Kadoodle. Cu uluitorul procent de 12%!

În interviul dat revistei *Science* în 2015[19], Epstein a explicat ce se întâmpla:

> Ne aşteptăm ca motorul de căutare să facă alegeri înţelepte. Participanţii la experiment îşi spun: „Ei bine, da, observ avantajarea unui anumit candidat, iar asta îmi transmite [...] că, de fapt, motorul de căutare îşi face datoria.

Dar, dacă luăm în considerare cât de multe dintre informaţiile noastre le primim de la algoritmi precum motoarele de căutare, este încă şi mai ameninţător să aflăm în ce măsură voluntarii credeau că îşi influenţaseră propriile opinii: „Când oamenii nu sunt conştienţi că sunt manipulaţi, ei tind să creadă că au adoptat o părere nouă din proprie voinţă", scria Epstein în lucrarea sa originală.[20]

Firește, Kadoodle nu este singurul algoritm acuzat că ar manipula, în mod subtil, opiniile politice ale cetăţenilor. Vom insista mai mult asupra acestui fapt în capitolul „Datele personale informatizate", însă, pentru moment, merită să reţinem ce sugerează experimentul acesta că simţim faţă de algoritmii care au dreptate în majoritatea timpului: sfârşim prin a crede că ei dau întotdeauna dovadă de un raţionament superior.[21] În scurt timp, nu mai suntem nici măcar conştienţi de propriul subiectivism faţă de aceştia.

Pretutindeni în jurul nostru, algoritmii oferă un fel de sursă convenabilă de autoritate. O cale uşoară de a delega responsabilitatea; o scurtătură pe care o întrebuinţăm fără să ne mai gândim. Cine, oare, va da clic de fiecare dată pe a doua pagină de rezultate de pe Google şi cine, oare, va analiza critic fiecare rezultat? Sau cine va trece pe la fiecare companie

aeriană să verifice dacă Skyscanner chiar afișează zborurile cele mai ieftine? Sau cine va scoate o riglă și o hartă rutieră ca să aibă confirmarea că GPS-ul chiar îi sugerează drumul cel mai scurt? Eu, una, nu, de asta puteți fi siguri!

Aici trebuie, însă, să facem o distincție. A avea încredere într-un algoritm fiabil de obicei este una, iar a avea încredere oarbă într-unul căruia nu-i înțelegi prea limpede caracteristicile este cu totul și cu totul altceva.

CÂND INTELIGENȚA ARTIFICIALĂ ÎNTÂLNEȘTE PROSTIA VERITABILĂ

În anul 2012, mai multe persoane cu dizabilități din Idaho au fost informate că le fusese redusă suma pe care o primiseră până atunci din partea programului federal de asistență medicală Medicaid.[22] Deși tuturor le fusese recunoscut dreptul de a beneficia de acest ajutor, statul le redusese susținerea financiară – fără nicio avertizare – chiar și cu 30%[23] lăsându-i pe bieții oameni să se chinuiască să-și achite costurile îngrijirilor medicale. Nu era vorba însă de o decizie politică, doar de consecința aplicării unui nou „instrument bugetar", recent adoptat de Idaho Department of Health and Welfare, ministerul sănătății din statul Idaho: un program de software care calcula automat cuantumul ajutorului financiar pe care trebuia să-l primească fiecare bolnav.[24]

Problema era că deciziile „instrumentului bugetar" nu păreau să aibă prea multă logică. Din câte își putea da seama cineva din afara sistemului, cifrele pe care le propunea programul erau, în mare parte, absolut arbitrare. Unor persoane li se dădeau bani mai mulți decât în anii precedenți, pe când altele și-au văzut bugetul redus cu zeci de mii de dolari, astfel încât riscau să fie nevoite să-și părăsească locuințele pentru a putea fi îngrijite într-o instituție publică.[25]

Incapabili să înțeleagă de ce le fuseseră reduse indemnizațiile sau să protesteze efectiv împotriva acestor reduceri, oamenii au apelat la American Civil Liberties Union (ACLU)* pentru a fi ajutați. Cazul lor a fost preluat de Richard Eppink, directorul juridic al filialei ACLU din Idaho[26], care a ținut să spună următoarele într-o postare pe blog din 2017:

> Am crezut, inițial, că acest caz avea să fie o chestie simplă și că tot ce urma să fim nevoiți să facem era să-i întrebăm pe reprezentanții statului: „În regulă, ce-ar fi să ne spuneți și nouă de ce au scăzut atât de mult sumele astea în dolari?"[27]

În realitate, a fost nevoie de patru ani, 4 000 de reclamanți și un proces colectiv pentru a se lămuri ce se întâmplase, de fapt.[28]

Eppink și echipa sa au început prin a cere amănunte cu privire la modul de funcționare a algoritmului, dar angajații Medicaid au refuzat să le explice cum calculaseră sumele respective, argumentând că software-ul care evalua cazurile constituia un „secret comercial", motiv pentru care nu putea fi făcut public.[29] Din fericire, judecătorul care prezida ședința nu a fost de acord cu acest lucru. Așa se face că instrumentul bugetar care exercita atâta putere asupra locuitorilor cu dizabilități a fost predat și s-a dovedit că nu era, nici pe departe, cine știa ce inteligență artificală sofisticată sau model matematic minunat construit, ci un simplu tabel în Excel![30]

În acesta, calculele ar fi trebuit să se bazeze pe istoricul unor cazuri, numai că informațiile erau atât de înțesate de *bug*-uri și erori, încât, în cea mai mare parte, erau absolut nefolositoare.[31] Lucru încă și mai grav, de îndată ce echipa de la

* Uniunea Americană pentru Libertăți Civile, ONG care luptă, conform propriei carte: „pentru apărarea și menținerea drepturilor individuale și a libertăților garantate tuturor cetățenilor acestei țări de către Constituția și legile Statelor Unite" (n. tr.).

ACLU a reușit să descâlcească ecuațiile, au fost descoperite „greșeli statistice fundamentale în modul în care era structurată însăși formula de calcul". Instrumentul bugetar produsese efectiv rezultate aleatorii pentru un număr imens de persoane! Algoritmul – dacă îl putem numi astfel – era de o calitate atât de proastă, încât, în cele din urmă, instanța l-a declarat neconstituțional.[32]

În acest caz s-au manifestat două șiruri paralele de erori umane. Mai întâi, cineva a scris un tabel în Excel numai bun de aruncat la gunoi; apoi, alte persoane au avut, cu naivitate, încredere în acesta! De fapt, „algoritmul" nu era decât munca de mântuială a cuiva, ambalată frumos sub forma unui cod. De ce erau, atunci, atât de dornici funcționarii care lucrau pentru stat să ia apărarea unui lucru atât de tâmpit?

Iată opinia lui Eppink în această chestiune:

> Ne-am confruntat cu acea slăbiciune pe care o avem cu toții când vine vorba de rezultatele obținute de un computer, pe care, pur și simplu, nu le mai punem sub semnul întrebării... Când un calculator generează ceva – care se bazează pe munca unui statistician care a analizat anumite informații și a inventat o formulă –, tindem să avem, orbește, încredere în formula respectivă, fără să ne mai întrebăm: „Hei, ia stai puțin! Cum funcționează, de fapt, chestia asta?"[33]

Știu că să descompui formule matematice pentru a analiza cum funcționează nu este tocmai distracția preferată a tuturor oamenilor (chiar dacă este a mea). Cu toate acestea, Eppink a scos în evidență o problemă incredibil de importantă cu privire la tendința noastră de a da crezare algoritmilor, de a-i lua de buni, fără să ne mai întrebăm cum funcționează, ce se petrece în spatele fațadei lor.

În anii în care, ca matematician, am lucrat cu informații și algoritmi, am ajuns să cred că singura modalitate de a aprecia cu obiectivitate dacă un algoritm este demn de

încredere este să mă lămuresc cum funcționează. Experiența pe care o dețin m-a învățat că algoritmii seamănă destul de mult cu scamatoriile unui magician. La început, par să fie, nici mai mult, nici mai puțin, decât magie pură, dar, de îndată ce ai aflat în ce constă scamatoria, misterul dispare. Adesea, în spatele fațadei se ascunde ceva amuzant de simplu (sau îngrijorător de neglijent, de nesăbuit). Așa că, în capitolele următoare și în legătură cu algoritmii pe care îi vom analiza, voi încerca, ori de câte ori îmi va fi cu putință, să vă fac să intuiți ce se petrece în spatele fațadei acestora. Atât cât să vedeți cum se fac scamatoriile informatice – chiar dacă nu îndeajuns de amănunțit, încât să le puteți realiza apoi voi înșivă.

Însă, chiar și pentru cei mai înrăiți pasionați de matematică, tot vor mai exista situații în care algoritmii le vor cere să se arunce, orbește, cu capul înainte și să-i creadă „pe cuvânt". Poate pentru că, așa cum este cazul cu Skyscanner sau cu rezultatele căutării pe Google, verificarea modului în care funcționează respectivele programe nu este posibilă. Sau poate pentru că, așa cum s-a întâmplat în cazul instrumentului bugetar din Idaho și în alte situații care vor fi prezentate în continuare, algoritmul respectiv este considerat un „secret comercial". Sau poate pentru că, așa cum se întâmplă cu anumite procedee de învățare automată, unei ființe umane îi este, pur și simplu, cu neputință să urmărească procesul logic care se desfășoară în interiorul algoritmului.

Așadar, vor exista momente în care vom fi siliți să cedăm controlul necunoscutului, chiar dacă vom fi pe deplin conștienți că algoritmul s-ar putea să greșească. Momente în care vom fi nevoiți să alegem între propriul raționament și cel al mașinii. Și când, dacă ne vom hotărî să avem, mai degrabă, încredere în instinctele noastre, nu în calculele sale, va trebui să dăm dovadă de foarte mult curaj.

Stanislav Petrov era un ofițer din armata rusă responsabil cu supravegherea sistemelor de avertizare nucleară timpurie care protejau spațiul aerian sovietic. Sarcina lui era să-și alerteze imediat superiorii în cazul în care computerul indica vreun semn că un atac american era iminent.[34]

La 26 septembrie 1983, Petrov era la datorie atunci când, după miezul nopții, sirenele au început să urle. Era semnalul de alarmă de care se temea toată lumea: sateliții sovietici identificaseră o rachetă inamică, una care se îndrepta către teritoriul Rusiei! Cum Războiul Rece era în toi, o lovitură armată era cât se poate de plauzibilă, numai că, totuși, ceva l-a cam pus pe gânduri pe Petrov și l-a făcut să nu ia nicio măsură. El nu era convins că avea deplină încredere în algoritm, căci acesta identificase numai cinci rachete, ceea ar fi constituit o lovitură de deschidere absurd de redusă pentru un atac nuclear american.[35]

Petrov a înlemnit în scaunul său. Din clipa aceea totul depindea de el: putea fie să tragă semnalul de alarmă și să arunce întreaga lume într-un război nuclear aproape sigur, fie să aștepte, ignorând uzanțele și știind că fiecare secundă care se scurgea însemna că, de fapt, conducătorilor țării sale le rămânea tot mai puțin timp la dispoziție pentru a lansa o contralovitură.

Din fericire pentru noi toți, Petrov a hotărât să aleagă cea de-a doua variantă. Nu putea fi absolut sigur că alarma se declanșase din greșeală, dar, după 23 de minute – care trebuie să i se fi părut, atunci, o eternitate –, când a fost limpede că nicio rachetă nucleară nu căzuse pe pământ rusesc, a știut că avusese dreptate. Algoritmul fusese acela care greșise.

Dacă sistemul ar fi acționat complet independent, fără să aibă nevoie de un om ca Petrov care să acționeze ca arbitru final, istoria s-ar fi desfășurat, fără îndoială, într-un mod diferit. Rusia ar fi lansat, aproape sigur, ceea ce ar fi considerat a fi un contraatac și ar fi declanșat, totodată, un război nuclear

în toată regula. Dacă putem să tragem o învățătură din toată povestea, atunci aceasta este că elementul uman trebuie să fie o parte indispensabilă a procesului: să dispui de o persoană cu drept de veto care să fie în măsură să verifice sugestiile făcute de un algoritm înainte de luarea unei decizii este singurul mod rațional de a evita greșelile.

În definitiv, numai oamenii vor simți că poartă povara răspunderii pentru deciziile pe care le iau. Un algoritm însărcinat cu transmiterea mesajelor până la nivelul oficialilor de la Kremlin nu ar fi așteptat nicio secundă să își ducă la îndeplinire însărcinarea, pentru că nu ar fi șovăit tot analizând consecințele posibile ale unei astfel de decizii. Pe de altă parte, ce a făcut Petrov? „Știam cât se putea de bine că nimeni nu ar fi reușit să-mi corecteze greșeala, dacă aș fi făcut-o."[36]

Singura problemă în legătură cu această concluzie este că nici pe ființele umane nu te poți baza chiar întotdeauna. Uneori, asemenea lui Petrov, vor avea dreptate să nu țină cont de sugetiile unui algoritm. Însă, adeseori, cel mai bine ar fi să ne ignorăm instinctele.

Ca să ofer un alt exemplu din domeniul securității, unde situațiile în care ființele umane ignoră în mod greșit un algoritm sunt, din fericire, rare, exact asta s-a întâmplat în cazul celebrei prăbușiri a montagnes russes-ului denumit The Smiler („Zâmbărețul"), din Alton Towers, cel mai mare parc de distracții tematic din Marea Britanie.[37]

În iunie 2015, doi ingineri au fost chemați să remedieze o defecțiune la un montagnes russes. După ce au rezolvat problema, au trimis un vagon gol pe circuit, să verifice dacă totul funcționa așa cum trebuie, însă nu au observat că acesta nu s-a mai întors. Din cine știe ce motiv, vagonul gol s-a răsturnat pe o pantă și s-a oprit la mijlocul circuitului.

Între timp, fără știrea inginerilor, personalul de deservire a adăugat la tren un vagon suplimentar pentru a face față cozilor din ce în ce mai mari. Odată ce au primit undă verde de la

cabina de control, angajații au început să umple vagoanele cu pasageri veseli, cărora le-au legat centurile, și au trimis primul tren pe traseu, fără să aibă habar de existența vagonului gol, înțepenit în mijlocul șinelor, care fusese trimis de ingineri și care stătea de acum chiar în drum.

Din fericire, arhitecții montagne russes-ului prevăzuseră o situație de acest gen, iar algoritmii de siguranță au funcționat exact așa cum fuseseră proiectați s-o facă. Pentru a evita o coliziune sigură, trenul ticsit de oameni a fost oprit automat în vârful primului urcuș, iar un semnal de alarmă s-a declanșat în cabina de control. Numai că inginerii – încrezători că abia reparaseră defecțiunea – au tras concluzia că sistemul automat de avertizare greșea.

Ignorarea algoritmului nu era o acțiune simplă: era necesar ca amândoi să fie de acord și să apese simultan câte un buton pentru a reporni montagne russes-ul. Procedând în felul acesta, trenul ticsit de oameni a fost trimis la vale și s-a lovit în plin de vagonul rămas pe șine. Consecințele au fost de-a dreptul groaznice. În urma impactului, mai multe persoane au suferit leziuni devastatoare și două adolescente și-au pierdut picioarele.

Aceste două situații de viață și de moarte, semnalele de alarmă declanșate în cazul lui Petrov și al montagnes russes-ului din parcul de distracții Alton Towers, ilustrează, în mod tragic, o dilemă profundă: în echilibrul de forțe dintre oameni și algoritmi, cine (sau ce) ar trebui să aibă ultimul cuvânt?

LUPTA PENTRU PUTERE

Această dezbatere durează de multă vreme. În 1954, Paul Meehl, profesor de psihologie clinică la Universitatea din Minnesota, a supărat o întreagă generație... umană atunci

când a publicat *Clinical versus Statistical Prediction**, o carte în care insista ferm asupra unuia din aspectele argumentației.[38]

În cartea sa, Meehl a comparat sistematic, riguros, performanțele oamenilor și ale algoritmilor într-o mare diversitate de situații (de la prezicerea notelor studenților și până la evaluarea sănătății mintale a pacienților) și a ajuns la concluzia că algoritmii matematici, oricât ar fi de simpli, vor face, aproape întotdeauna, previziuni mai bune decât oamenii.

Nenumărate alte studii, efectuate în jumătatea de secol care a urmat, au confirmat constatările lui Meehl. Dacă sarcina pe care o ai de îndeplinit presupune efectuarea oricărui tip de calcul, poți să pariezi de fiecare dată că algoritmii se vor descurca mai bine: de la stabilirea diagnosticelor medicale sau a vânzărilor prognozate, la prezicerea tentativelor de suicid sau a gradului de satisfacție în carieră, precum și atunci când trebuie evaluat absolut orice, de la condiția fizică necesară pentru îndeplinirea stagiului militar și până la preconizarea performanțelor ca student.[39] Mașina poate că nu va acționa perfect, dar a-i lăsa omului drept de veto asupra algoritmilor nu va duce, în anumite cazuri, decât la înmulțirea erorilor.**

Poate că asta nu ar trebui să ne surprindă. Noi nu am fost concepuți pentru a face calcule. În fond, când mergem la supermarket, nu dăm peste un șir de casieri care ne verifică din priviri cumpărăturile făcute, pentru a estima cam cât ar costa. Pentru asta avem un algoritm (incredibil de simplu)

* „Decizia clinică împotriva celei statistice" (n. tr.).

** În mod surprinzător, una dintre rarele excepții în ceea ce privește superioritatea algoritmilor în a face previziuni provine dintr-o serie de studii realizate la sfârșitul anilor 1950 și pe parcursul anilor 1960, despre „diagnosticarea" (termenul le aparține cercetătorilor vremii, nicidecum mie!) homosexualității. În acele exemple, raționamentul uman a făcut predicții mult mai bune, depășind ca performanță tot ce reușea să prevadă algoritmul – ceea ce sugerează că anumite lucruri sunt atât de specifice ființei umane, încât informațiile și formulele matematice vor avea întotdeauna dificultăți în a le descrie (n. a.).

care socoteşte în locul nostru. Şi, în majoritatea timpului, ar fi mai bine să lăsăm maşina să calculeze. Este exact ca în gluma aceea care circulă în lumea piloţilor de avion, potrivit căreia, de fapt, cel mai bun echipaj de zbor este compus din trei membri: un pilot, un computer şi un câine. Calculatorul face avionul să zboare, iar pilotul hrăneşte câinele. Patrupedul se află acolo ca să-l muşte pe om dacă acesta ar încerca să se atingă de computer.

Numai că în relaţia noastră cu maşinile se manifestă un paradox. Deşi tindem să acordăm o încredere nețărmurită tuturor lucrurilor pe care nu le înţelegem, de îndată ce ştim că un algoritm poate să facă greşeli, avem obiceiul destul de enervant de a reacţiona exagerat şi de a-l înlătura cu totul, revenind la propriul nostru raţionament deficitar. Cercetătorii cunosc acest comportament, pe care l-au denumit „aversiune faţă de algoritmi". Oamenii au mai puţină toleranţă faţă de erorile unui algoritm decât faţă de propriile greşeli, chiar dacă acestea din urmă sunt mai mari, mai grave.

Fenomenul a fost demonstrat în repetate rânduri în cadrul unor experimente[40] şi, într-o oarecare măsură, îl puteţi recunoaşte şi în propriile dumneavoastră manifestări. Spre exemplu, ori de câte ori Citymapper mă anunţă că, de fapt, călătoria mea va dura mai mult decât m-am aşteptat, nu mă pot abţine să cred că tot eu ştiu mai bine (chiar dacă, de cele mai multe ori, asta înseamnă că voi sfârşi prin a întârzia). Cu toţii am făcut-o pe Siri „idioată"* măcar o dată, uitând cumva, pe parcurs, cât de uimitoare sunt realizările tehnologice care au dus la alcătuirea unui asistent personal vorbitor pe care să-l poţi ţine în mână! În perioada de început a utilizării aplicaţiei cu GPS Waze pe mobil, m-am trezit, la un moment dat, prinsă într-un ambuteiaj după ce mă convinsesem singură că urma

* Deoarece majoritatea utilizatorilor români folosesc varianta americană a acestei aplicaţii, a cărei voce este aceea a actriţei Susan Bennett, am optat ca aici şi în continuare să o considerăm pe Siri de sex feminin (n. red.).

s-ajung mai repede la destinație dacă aveam să o iau pe niște străduțe lăturalnice, nu pe ruta pe care mi-o indicase Waze. (De-atunci am aflat că niciodată nu aveam dreptate când presupuneam asta.) Acum am ajuns să am încredere în Waze și – asemenea lui Robert Jones și a BMW-ului său – o voi urma orbește, oriunde m-ar conduce (deși cred, totuși, că aș refuza să mă arunc într-o prăpastie).

Tendința aceasta de a vedea lucrurile în alb și negru – de a privi algoritmii fie ca niște stăpâni omnipotenți, fie ca pe o grămadă inutilă de gunoaie – este, întrucâtva, o problemă în era noastră high-tech. Dacă ne propunem să profităm la maximum de tehnologie, trebuie să găsim o cale de a fi obiectivi. Și ar fi bine să învățăm din greșeala lui Kasparov și să ne recunoaștem greșelile, să punem sub semnul întrebării reacțiile instinctive și să fim ceva mai conștienți de sentimentele noastre față de algoritmii care au ajuns să ne înconjoare. Pe de altă parte, ar trebui să coborâm algoritmii de pe piedestal, să-i studiem cu ceva mai multă atenție și să ne întrebăm dacă sunt, cu adevărat, în stare de tot ceea ce pretind c-ar putea face. Este singura modalitate de a decide dacă merită să dețină toată puterea care le-a fost conferită.

Din nefericire, toate acestea sunt, de multe ori, mai ușor de zis decât de făcut. Deseori nu prea vom avea vreun cuvânt de spus în legătură cu puterea și răspândirea algoritmilor, chiar și atunci când vine vorba despre cei care ne afectează în mod direct.

Asta este, mai cu seamă, adevărat pentru algoritmii care fac schimb cu cel mai modern produs de bază: informația. Algoritmii care ne urmăresc în tăcere pe internet, cei care culeg informațiile noastre personale, ne invadează intimitatea și ne lasă impresia că am avea mână liberă, numai pentru a ne influența apoi, în mod subtil, comportamentul. Iar în această adevărată „furtună perfectă" a încrederii, puterii și influenței prost plasate, consecințele au potențialul de a modifica în mod fundamental societatea în care trăim.

Capitolul 2

DATELE PERSONALE INFORMATIZATE

În 2004, la puțin timp după ce crease Facebook, studentul Mark Zuckerberg a avut următorul schimb de mesaje instantanee, pe chat, cu un prieten:

ZUCK: Mda, așa că, dac-ai vreodată nevoie de informații despre oricine de la Harvard...

ZUCK: Nu trebuie decât să-mi ceri!

ZUCK: Am peste 4 000 de e-mail-uri, poze, adrese...

[NUMELE PRIETENULUI ESTE CENZURAT]: Ce? Cum ai reușit una ca asta?

ZUCK: Oamenii mi le-au comunicat de bunăvoie.

ZUCK: Habar n-am de ce.

ZUCK: Cic-ar avea „încredere în mine"!

ZUCK: Tâmpiți ordinari![1]

Ca urmare a scandalului în care a fost implicat Facebook în 2018, aceste cuvinte au fost publicate în mod repetat de jurnaliști, din dorința de a insinua că ar fi existat o atitudine machiavelică față de confidențialitatea informațiilor în cadrul companiei. Eu, una, cred c-am putea da dovadă de mai multă mărinimie atunci când vine vorba să interpretăm comentariile pline de lăudăroșenie ale unui tânăr de 19 ani. Dar mai cred și că Zuckerberg se înșela. Oamenii nu îi oferiseră, pur și simplu,

informații despre ei înșiși. Ci i le puseseră la dispoziție ca parte a unui schimb. Pentru ele, li se oferea acces la un algoritm care le permitea să țină legătura, fără restricții, cu prietenii și familia. În plus, li se punea la dispoziție un loc în care să-și împărtășească experiențele de viață cu alte persoane. Era rețeaua lor privată în imensitatea World Wide Web-ului. Nu știu ce părere aveți, dar, pe vremea aceea, eu, una, în mod cert consideram că era un schimb echitabil.

Există, însă, o singură fisură în raționamentul acesta: nu suntem întotdeauna conștienți de implicațiile pe termen mai lung ale unui asemenea schimb. Ne este, arareori, limpede ce pot face informațiile noastre sau cât de valoroase pot deveni ele atunci când sunt introduse într-un algoritm inteligent. Și, în consecință, nu ne este clar nici cât de ieftin am fost cumpărați.

PUȚIN CÂTE PUȚIN...

Companiile care gestionează supermarketuri au fost printre primele care au recunoscut valoarea strângerii informațiilor despre o persoană. Într-un domeniu în care firmele concurează permanent pentru a atrage atenția cumpărătorilor (cu scopul de a adăuga mici preferințe în comportamentul lor la shopping* care le vor întări, în timp, loialitatea față de marca produsului, față de brand), fiecare ușoară îmbunătățire contează, căci, treptat, se adaugă altora similare și se poate transforma într-un avantaj uriaș. Acesta a fost motivul care a dus

* *Buying behaviour* (engl.), „comportament de/la cumpărături/shopping". Am optat să păstrăm și anglicismul shopping în traducerea acestei sintagme, deoarece, conform *Dicționarului explicativ al limbii române* (DEX, ediția din 2009), el este cuprinzător, descriind mai multe lucruri, nu doar simpla achiziționare a unor produse, căci este definit drept: „vizitarea magazinelor pentru a vedea vitrinele, galantarele și a face cumpărături" (n. red.).

la organizarea experimentului revoluționar din 1993 realizat de supermarketul britanic Tesco.

Sub îndrumarea echipei formate din soții Edwina Dunn și Clive Humby și lansat în magazine anume selectate, Tesco a emis noul său Clubcard – un card plastifiat de dimensiunea și forma unui card de credit, pe care clienții îl puteau prezenta la casă atunci când își achitau cumpărăturile. Schimbul era simplu: la fiecare tranzacție în care folosea Clubcard, clientul acumula puncte pe care le putea utiliza pentru viitoare cumpărături din magazin, câtă vreme Tesco avea se înregistreze acea vânzarea și s-o asocieze cu numele clientului.[2]

Informațiile culese în acel prim experiment cu ajutorul lui Clubcard au fost extrem de limitate. Pe lângă numele și adresa clientului, programul nu înregistra decât ce sumă cheltuise acesta și când anume, nu și ce produse avusese în coșul de cumpărături. Totuși, din relativ modesta recoltă de informații, Dunn și Humby au reușit să extragă niște concluzii extrem de valoroase.

Ei au descoperit că un număr mic de clienți fideli realiza un volum mare din vânzări. Au aflat, examinând cod poștal după cod poștal, ce distanță erau oamenii dispuși să parcurgă până la magazine. Au văzut care erau cartierele în care concurența avea mai mult succes și cartierele în care avea preponderență Tesco. Informațiile le dezvăluiau ce clienți reveneau în fiecare zi și care își amânau cumpărăturile pentru sfârșitul săptămânii. Înarmați cu aceste cunoștințe, au trecut la treabă: au încercat să transforme comportamentul la shopping al clienților care utilizau Clubcard, trimițându-le, prin poștă, diverse cupoane. Clienților care cheltuiau mai mult li s-au oferit vouchere cu sume cuprinse între 3 și 30 de lire sterline. Celor care cheltuiau sume mici, li s-a trimis un stimulent mai modest – între 1 și 10 lire sterline. Iar rezultatele au fost uimitoare. Aproape 70% dintre cupoane au fost răscumpărate, iar în magazine clienții și-au ticsit coșurile: persoanele care dețineau un Clubcard au

cheltuit, per ansamblu, cu 4% mai mult decât cele care nu deţineau unul.

Pe 22 noiembrie 1994, Clive Humby a prezentat concluziile experimentului său membrilor consiliului de administraţie al companiei Tesco. Le-a arătat toate informaţiile culese, ratele de răspuns, dovezile satisfacţiei clienţilor, creşterile explozive ale vânzărilor. Membrii consiliului au ascultat în tăcere. La sfârşitul prezentării, preşedintele de atunci al firmei a vorbit primul:

> Ceea ce mă sperie în toată chestiunea asta, a spus el, este că voi aţi aflat mai multe informaţii despre clienţii mei în trei luni, decât am aflat eu în 30 de ani![3]

Clubcard-ul a fost extins la toţi clienţii Tesco şi i se recunoaşte, pe scară largă, meritul de a fi ajutat compania să o ia înaintea principalului său rival, Sainsbury's, şi să devină cel mai mare lanţ de supermarketuri din Marea Britanie. Cu timpul, informaţiile colectate au devenit mai detaliate, ceea ce a făcut ca obiceiurile de shopping ale clienţilor să devină tot mai uşor de urmărit.

Chiar de la începuturile cumpărăturilor on-line, echipa* a introdus elementul cunoscut drept *My Favourites***, în care orice obiect cumpărat cu ajutorul cardului de fidelitate apărea evidenţiat atunci când clientul se autentifica pe site-ul Tesco. Asemenea Clubcard-ului, şi acest element a avut un succes răsunător. Oamenii puteau găsi cu uşurinţă produsele pe care le doreau, fără să mai navigheze printr-o mulţime de pagini. Vânzările au crescut, iar clienţii au fost mulţumiţi.

* Dunnhumby, numele companiei de servicii de gestionare a datelor clienţilor (*customer data services*, engl.) înfiinţate de soţii Edwina Dunn şi Clive Humby în 1989, în propria bucătărie, care a ajuns în prezent să aibă sucursale în peste 30 de ţări şi să fie estimată a avea o valoare cuprinsă între unu şi trei miliarde de dolari americani (n. red.).

** „Preferinţele mele" (engl.) (n. tr.).

Nu toți, însă. La puțin timp după lansarea acestui element, o femeie a contactat compania Tesco și s-a plâns că informațiile care i se ofereau ei erau greșite. Făcuse cumpărături on-line și, printre produsele ei din lista *My Favourites* observase și prezervative! Nu puteau fi ale soțului ei, explicase clienta, pentru că el nu folosea așa ceva. La această solicitare, analiștii de la Tesco au studiat informațiile și au descoperit că lista era corectă. Cu toate acestea, în loc să fie cauza unei rupturi conjugale, au luat decizia diplomatică de a-și cere scuze pentru „informațiile incorecte" și de a șterge obiectele ofensatoare din lista de produse favorite a clientei.

După cum a scris Clive Humby în cartea lui despre Tesco, procedura a devenit acum politica neoficială a companiei. Astfel, ori de câte ori în datele informatizate ale unui client apare ceva care dezvăluie cam prea multe, cei de la Tesco își cer scuze și șterg informațiile. Este o atitudine pe care și-a însușit-o și Eric Schmidt, care, pe când era președintele executiv al Google, spunea că încerca să gândească în termenii unei limite imaginare care ar despărți lucrurile acceptabile de acelea sinistre (*creepy line*):

> Politica Google este de a se apropia cât mai mult de granița care ne desparte de chestiile de-a dreptul sinistre, fără să o depășească, însă.[4]

Totuși, dacă vei culege suficiente informații, e greu de spus ce vei descoperi. Deoarece cumpărăturile nu reprezintă numai ceea ce consumi. Ele sunt ceva profund personal. Dacă vei studia cu atenție comportamentul la shopping al cuiva, el îți va dezvălui, de cele mai multe ori, tot felul de amănunte despre cine este acesta ca persoană. Uneori – cum a fost în cazul prezervativelor –, va divulga lucruri pe care, poate, ar fi mai bine să nu le știi. Însă, cel mai adesea, în tot acest noian

de informații, se ascund frânturile unor descoperiri uimitoare care pot fi folosite în beneficiul unei companii.

MAGAZINUL TARGET

În anul 2002, Target, compania de supermarketuri americane specializate în comercializarea produselor cu reduceri, a început să caute pattern-uri mai neobișnuite în informațiile colectate.[5] Target vinde orice, de la lapte și banane, la jucării de pluș sau mobilă de grădină, și – aproape cam ca oricare alt comerciant cu amănuntul de la începutul secolului al XXI-lea – are modalitățile lui de a utiliza cifrele de pe cardurile de credit și răspunsurile la sondaje pentru a face legătura între clienți și tot ceea ce au achiziționat aceștia vreodată din magazin, ceea ce le permite să analizeze ce cumpără oamenii.

O știre (care nu mai trebuie să le fie repetată cititorilor americani) a făcut de rușine compania în Statele Unite: se pare că, la un moment dat, analiștii de la Target și-au dat seama că o creștere bruscă a achiziționării de către femei a loțiunilor de corp neparfumate indica, adesea, că respectivele cliente urmau să se înscrie în registrul magazinului* pentru ca rudele și prietenii lor să poată cumpăra articolele pe care și-ar fi dorit să le primească în dar la *baby-shower*** În datele informatizate fusese descoperit un indiciu plin de semnificații: când femeile intrau în al doilea trimestru de sarcină și începeau să își facă griji din cauza vergeturilor, cumpărarea de creme hidratante pentru menținerea suppleței pielii era un indiciu care prevestea ceea ce avea să

* În SUA chiar și supermarketurile le permit clienților să își creeze o listă cu lucrurile aflate la vânzare în respectiva unitate comercială pe care ar dori să le primească drept cadou cu diverse ocazii (nuntă, botezul copiilor etc.). O astfel de listă poartă numele de *gift registry* (engl., „registru al cadourilor") (n. red.).

** Petrecere organizată în timpul sarcinii de către viitoarea mamă sau de către prietenii ori rudele acesteia, pentru a celebra apropiata sosire pe lume a copilului (n. tr.).

urmeze. Dacă ai fi derulat înapoi istoricul cumpărăturilor, ai fi observat că aceleași femei trecuseră pe la Target și se aprovizionaseră cu diverse vitamine și suplimente alimentare, cum ar fi calciu și zinc. Dac-ai fi derulat apoi înainte aceeași listă de cumpărături, informațiile ți-ar fi sugerat chiar și data la care se născuse bebelușul – indicată de faptul că după aceea, femeia cumpăra din magazin pungi foarte mari de vată.[6]

Mamele însărcinate sunt visul oricărui comerciant cu amănuntul. Dacă ți-ai asigurat fidelitatea lor cât timp erau însărcinate, sunt șanse mari ca ele să continue să-ți folosească produsele mult timp după nașterea copilului. În definitiv, obiceiurile la shopping se formează iute atunci când un bebeluș flămând care urlă din toți rărunchii îți solicită atenția în timpul cumpărăturilor săptămânale. Informațiile de acest gen erau de o inestimabilă valoare, căci îi puteau oferi companiei Target un avantaj față de rivalele sale în atragerea cumpărăturilor pe care urma să le facă mama respectivă.

De aici încolo, lucrurile au fost simple. Target a rulat un algoritm care avea să ofere un punctaj clientelor, în funcție de probabilitatea ca acestea să fie însărcinate. Dacă probabilitatea depășea un anumit prag, magazinul avea că îi trimită în mod automat femeii în cauză o serie de cupoane cu reduceri pentru tot soiul de lucruri pe care aceasta le-ar fi putut considera utile: scutece, loțiuni, șervețele umede pentru copii și așa mai departe.

Amintita practică era, pe atunci, necontroversată. Apoi, cam la un an după introducerea acestui instrument, tatăl unei adolescente a intrat ca o furtună într-un magazin Target din Minneapolis, cerând să discute cu responsabilul unității. Fiicei sale îi fuseseră trimise cupoane cu produse pentru femei gravide, iar el era revoltat că magazinul părea să considere firești sarcinile în rândul adolescentelor. Responsabilul și-a cerut mii de scuze, iar câteva zile mai târziu a sunat la

domiciliul bărbatului pentru a-şi exprima încă o dată părerea de rău pentru toată încurcătura. Numai că, între timp, conform relatării din *New York Times*, tatăl s-a trezit, la rândul său, că avea de prezentat nişte scuze.

„Am avut o discuţie cu fiica mea", i-a spus el responsabilului. „Se pare că în casa mea s-au petrecut unele lucruri de care nu prea am avut habar. Fiica mea urmează să nască în august."

Nu ştiu ce părere aveţi, dar un algoritm care va înştiinţa un părinte că fiica sa este însărcinată înainte ca el să aibă prilejul să afle personal acest lucru, mie mi se pare că trece mult dincolo de graniţa care ne desparte de chestiile de-a dreptul sinistre. Această situaţie stânjenitoare nu a fost, însă, suficientă pentru a convinge firma să renunţe definitiv la respectiva strategie de marketing.

Un director al companiei Target a explicat următoarele:

> Am aflat că, atât timp cât o femeie însărcinată nu crede că a fost urmărită, ea va folosi cupoanele noastre de reduceri. Va presupune numai că toţi cei din blocul în care locuieşte au primit acelaşi pliant pentru scutece şi pătuţuri. Câtă vreme nu o speriem, metoda funcţionează.

Aşadar, Target încă are un algoritm capabil să anticipeze sarcinile care rulează în fundal – aşa cum procedează, în prezent, majoritatea comercianţilor. Singura diferenţă este că, în prezent, când va trimite cupoane cu reduceri pentru diversele produse dedicate femeilor însărcinate, le va amesteca printre altele pentru articole de uz general, astfel încât clientele să nu observe că au fost vizate în mod direct. O reclamă la pătuţuri pentru nou-născuţi ar putea să apară lângă una pentru pahare de vin. Sau un cupon cu reduceri pentru hăinuţe destinate bebeluşilor va fi postată pe site-ul companiei lângă o reclamă la o marcă de apă de colonie.

Target nu este singura firmă care folosește aceste metode. Relatările despre ceea ce se poate deduce din datele personale ale cuiva ajung rareori în presă, însă algoritmii sunt prezenți, ascunzându-se, tăcuți, în spatele fațadelor corporatiste. Cu aproximativ un an în urmă, am stat de vorbă cu analistul-șef* al unei companii care vindea asigurări. Prin intermediul unei strategii de marketing de loializare** a clienților unui supermarket, el și colegii lui aveau acces la toate detaliile privind obiceiurile de shopping ale acestora. Analizând informațiile, analiștii au descoperit că toți cei care găteau acasă erau mai puțin înclinați să emită cereri de despăgubire, deși locuințele le erau asigurate, fiind, prin urmare, clienți mai profitabili pentru compania sa. Acesta este un exemplu de descoperire care pare de bun-simț, în mod intuitiv. Probabil că nu există prea multe puncte comune între cei care sunt dispuși să investească timp, efort și bani pentru a pregăti de la zero diverse feluri de mâncare complicate și cei care își lasă copiii să joace fotbal în casă. Dar de unde au știut analiștii companiei de asigurări care cumpărători găteau acasă și care nu? Ei bine, pur și simplu, prezența unor anumite articole în coșul de cumpărături al unor persoane corespundea cu o rată redusă de cereri de despăgubire. Cel mai important articol, mi-a declarat analistul-șef, acel produs care dezvăluie mai mult decât orice altceva că respectivul cumpărător este o persoană responsabilă și mândră de locuința sa, e... feniculul proaspăt.

* *Chief data officer* (engl., CDO), director al unei corporații a cărui responsabilitate o constituie gestionarea și utilizarea datelor informatizate prin diverse mijloace: colectare, prelucrare, analizare, *data mining*, schimb de informații ș.a. (n. red.).

** *Loyalty scheme* (engl.), strategie de marketing care își propune să îi transforme pe clienții deja existenți în cumpărători dedicați, care nu mai folosesc alte magazine. Deși termenul nu este atestat de dicționarele românești, am optat să îl păstrăm deoarece este larg folosit în lucrările de marketing, alături de „fidelizare" (n. red.).

Dacă poți să deduci una ca asta din comportamentul la shopping al persoanelor în lumea reală, imaginați-vă ce se poate deduce dacă ai acces la mai multe informații. Imaginați-vă numai cât de multe lucruri se pot afla despre cineva dacă dispui de evidența întregii sale activități on-line!

VESTUL SĂLBATIC

Palantir Technologies este una dintre cele mai de succes companii abia înființate din Sillicon Valley. A fost întemeiată în 2003 de Peter Thiel (ultracelebrul cofondator al PayPal), iar, la cea mai recentă evaluare disponibilă înainte de publicarea acestei cărți, s-a estimat că ar valora uluitoarea sumă de 20 de miliarde de dolari.[7] Are cam aceeași valoare de piață ca și Twitter, deși sunt șanse să nu fi auzit niciodată de acei de la Palantir. Totuși – și vă rog să mă credeți când vă spun asta! – Palantir a auzit, mai mult ca sigur, de fiecare dintre noi.

Palantir este face parte dintr-o nouă specie de companii, cunoscute drept „brokeri de date"*, care cumpără și colectează informațiile personale ale oamenilor, pentru ca apoi să le revândă sau să le împartă cu alte firme pentru a obține profit. Și există multe alte companii asemănătoare: Acxiom, Corelogic, Datalogix, eBureau – o grămadă de firme uriașe cu care probabil că nu ați interacționat niciodată în mod direct, dar care vă monitorizează și vă analizează neîncetat comportamentul.[8]

De fiecare dată când faceți cumpărături on-line, de fiecare dată când vă înscrieți la un newsletter, vă înregistrați pe un site, vă interesați de o mașină nouă, completați un certificat de garanție, cumpărați o locuință nouă sau vă înregistrați să votați – adică de fiecare dată când oferiți cuiva orice fel de date personale –, toate aceste informații sunt adunate și vândute apoi unui broker de date. Vă aduceți aminte când i-ați descris unui

* *Data broker* (engl.) (pentru definiție, vezi glosar, p. 316) (n. red.).

agent imobiliar ce fel de locuință căutați? Ei bine, informațiile acelea i-au fost deja vândute unui broker de date. Mai știți toate amănuntele acelea pe care le-ați tastat la un moment dat pe un website care compara prețurile diferitelor polițe de asigurări? Și ele i-au fost deja vândute unui broker de date. În anumite situații, chiar și întregul istoric al browser-ului vostru poate fi compactat și vândut.[9]

Cade în sarcina brokerului să combine toate aceste informații, să coreleze diferitele fragmente din toate frânturile digitale pe care le-a cumpărat și le-a strâns, iar apoi să creeze un singur fișier detaliat despre fiecare dintre noi: un profil psihologic alcătuit din date, urma ta digitală, „umbra" ta informatică. În sensul cel mai literal, în unele dintre aceste baze de date ale brokerilor de informații ai putea să deschizi un fișier digital purtând numărul tău de identificare (un număr care ție nu-ți va fi comunicat niciodată) care conține urmele a tot ceea ce ai făcut vreodată: nume, data nașterii, credință religioasă, obiceiuri de vacanță, mod de utilizare a cardului de credit, avere, greutate, înălțime, afinități politice, ce jocuri obișnuiești să joci, ce dizabilități ai, ce medicamente folosești, dacă ai făcut vreodată vreun avort, dacă părinții tăi sunt divorțați sau nu, dacă ai o predispoziție spre a deveni dependent cu ușurință, dacă ai fost victima unui viol, ce părere ai despre interzicerea armelor de foc, care este orientarea sexuală pe care o afișezi și care e adevărata ta orientare sexuală sau cât de credul ești... Practic, pentru fiecare dintre noi există mii și mii de detalii, incluse în mii și mii de categorii, în fișiere stocate pe niște servere aflate cine știe pe unde.[10]

Asemenea predicțiilor făcute de analiștii companiei Target despre sarcinile clientelor, multe dintre aceste informații pot fi deduse cu ușurință. Un abonament la revista *Wired* ar putea însemna că ești interesat de tehnologie; un permis de portarmă – că te pasionează vânătoarea. Pe tot acest parcurs, brokerii folosesc algoritmi inteligenți, dar simpli, pentru

a spori mereu cantitatea de informații pe care o dețin. Este exact ceea ce făceau și supermarketurile, dar pe o scară cu mult mai mare.

Și, datorită acestor profiluri, avem parte și de o mulțime de beneficii. Astfel, brokerii de date își folosesc toate cunoștințele pe care le dețin despre cine suntem și cum ne comportăm on-line pentru a-i împiedica pe escroci să se dea drept consumatori nevinovați. De asemenea, tot pentru că ni se cunosc preferințele, dar și lucrurile care ne displac, reclamele care ne sunt prezentate în timp ce navigăm pe internet sunt pe cât de relevante e cu putință pentru interesele și nevoile noastre. Aceasta aproape că ne asigură o experiență mai plăcută decât aceea în care am fi agresați zi de zi de reclame transmise la grămadă, tuturor,* la vreo asigurare de credit oarecare ori la serviciile oferite de cine știe ce avocat specializat în procesele deschise de aceia care pretind c-ar fi fost răniți de diverse persoane. Mai mult decât atât, pentru că mesajele pot fi țintite direct spre consumatorii potriviți, care sunt interesați de ele, publicitatea devine mai ieftină, în general, astfel încât firmele mici care oferă produse de excepție pot ajunge la noi nișe de piață, la noi categorii de consumatori, ceea ce este bine pentru toată lumea.

Însă, așa cum sunt convinsă că v-ați dat seama deja, pot apărea și o mulțime de probleme în momentul în care începi să ne separi pe noi, ca indivizi, într-o serie de categorii. Voi trece la acest subiect în curând, dar, mai întâi, cred că merită să explic, pe scurt, procesul nevăzut care are loc pentru ca o reclamă on-line să ajungă la noi atunci când dăm clicuri pe diverse site-uri de pe net, precum și rolul pe care-l joacă brokerii de date în acest proces.

Să ne imaginăm, așadar, că aș deține o agenție de voiaje de lux cu numele fictiv de Fry's. În decursul anilor, am reușit să fac oamenii să se înregistreze pe site-ul companiei mele

* *Mass market adverts* (engl.), „[campanie de] reclame care face parte dintr-o strategie de tip mass market" (vezi glosar, p. 318) (n. red.).

dacă au manifestat interes față de el, iar de acum am o listă cu adresele lor de e-mail. Dacă aș vrea să aflu mai multe despre utilizatorii mei – de exemplu, ce fel de vacanțe preferă –, aș putea să trimit lista cu adresele lor de e-mail unui broker de date, care le-ar căuta numele în sistem și mi-ar înapoia lista, completată însă cu diverse informații relevante. Este ca și când s-ar adăuga o coloană nouă într-un tabel. De acum, ori de câte ori un utilizator va vizita site-ul Fry's, eu voi vedea că respectivul are o anumită preferință pentru insulele tropicale și, astfel, o să-i ofer o reclamă la o excursie în Hawaii.

Aceasta ar fi prima mea opțiune. Pentru cea de-a doua, haideți să ne imaginăm că Fry's are ceva spațiu suplimentar pe site, pe care sunt dispusă să-l vând altor agenții de publicitate. Voi contacta din nou un broker de date și-i voi da informațiile pe care le dețin despre utilizatorii mei. Brokerul va căuta companii care să-și dorească să-și prezinte reclamele pe site-ul agenției mele. Și, doar de dragul de a continua această povestioară, haideți să ne imaginăm că o companie care vinde o loțiune de protecție solară și bronzare este deosebit de interesată să-i facă reclamă. Ca să o conving că Fry's deține audiența căreia vânzătorul de cosmetice respectiv ar dori să i se adreseze, brokerul ar putea să-i dezvăluie unele trăsături ale utilizatorilor site-ului meu pe care le-a dedus analizându-le datele informatizate: spre exemplu, care este procentul utilizatorilor mei care au păr roșcat ori ceva de genul acesta. Sau vânzătorul loțiunii ar putea să-i ofere propria listă cu adresele de e-mail ale utilizatorilor săi, iar brokerul și-ar putea da seama exact câte puncte comune există între cele două segmente de public. Dacă vânzătorul loțiunii de protecție solară este de acord, reclamele lui vor apărea pe site-ul Fry's – și atât brokerul, cât și eu, vom fi plătiți.

Până acum, metodele acestea nu depășesc cu mult strategiile pe care specialiștii în marketing le-au utilizat dintotdeauna pentru a-și viza clienții. Dar deja, când trecem

la cea de-a treia opțiune, lucrurile încep să devină cam prea sinistre. De data aceasta, Fry's caută niște clienți noi și speră să atragă bărbați și femei care au peste 65 de ani, cărora le plac insulele tropicale și dispun de venituri mari, nădăjduind că aceștia vor dori să meargă în vreuna dintre noile croaziere de lux în Caraibe pe care le oferă. Așa că, în calitate de patron, voi contacta un broker de informații, care va căuta în baza lui de date și-mi va găsi o listă cu diverse persoane care se potrivesc acestei descrieri.

Să ne imaginăm, așadar, că numele dumneavoastră se află pe lista aceea. Brokerul nu-i va dezvălui, în niciun caz, vreo identitate companiei Fry's. Însă va ști ce alte site-uri vizitați cu regularitate. Sunt șanse ca și brokerul să aibă o relație de afaceri cu unul dintre deținătorii site-urile voastre preferate. Poate că e vorba de un site de socializare, poate că e vorba de unul de știri – ceva de genul acesta. Imediat ce v-ați autentificat, fără să bănuiți nimic, pe unul dintre site-urile preferate, brokerul va primi o înștiințare (*ping*) care îl anunță că navigați pe una dintre paginile acestuia. Practic instantaneu, brokerul va răspunde prin plasarea unui „steguleț" – cunoscut drept *cookie*– în calculatorul tău. Această *cookie** acționează ca un semnal pentru toate tipurile de site-uri de pe internet, transmițându-le că sunteți o persoană căreia ar trebui să i se prezinte o reclamă la croazierele în Caraibe ale companiei Fry's. Așa că, fie că vă doriți acest lucru, fie că nu, reclamele vă vor urmări pretutindeni, oriunde veți naviga pe internet.

* Reclamele nu sunt singurul motiv pentru care există *cookie*-uri, căci acestea mai sunt folosite de site-uri ca să vadă dacă sunteți autentificat sau nu (și să știe dacă este sigur să trimită pe această cale informații cu caracter confidențial) și ca să vadă dacă sunteți un vizitator care revine pe o anumită pagină (pentru a putea obține o creștere a prețurilor pe site-ul unei companii aeriene, de exemplu, sau pentru a vă trimite prin e-mail un cod pentru obținerea unei reduceri la un magazin virtual de îmbrăcăminte) (n. a.).

Şi aici apare prima problemă. Ce se întâmplă dacă nu-ţi doreşti să vezi acea reclamă? Este adevărat, să fii bombardat cu imagini ale unor croaziere în Caraibe nu poate fi decât un inconvenient minor, dar există şi reclame care chiar pot afecta profund o persoană.

Când Heidi Waterhouse a pierdut o sarcină pe care şi-o dorise nespus de mult[11], ea s-a dezabonat de la toate e-mailurile săptămânale care o ţineau la curent cu dezvoltarea bebeluşului, spunându-i de mărimea cărui fruct ajunsese fătul într-o anumită etapă. Şi-a anulat, totodată, şi abonamentele la toate listele de corespondenţă, precum şi toate listele de lucruri dorite* la care se înscrisese în aşteptarea plină de înfrigurare a momentului naşterii. Numai că, aşa cum avea să le relateze ea unui grup de dezvoltatori la o conferinţă din 2018, nicio putere din lume nu o putea ajuta să îşi anuleze „abonarea" la puzderia de reclame pentru femeile însărcinate care o urmăreau pretutindeni pe internet. Urmele digitale ale sarcinii au continuat să circule singure, ca nişte umbre informatice, deşi, de acum, nu mai existau nicio mamă şi niciun bebeluş. „Niciunul dintre cei care au creat sistemul nu s-a gândit vreodată la astfel de consecinţe", a explicat ea.

Fiindcă este un sistem care, fie din cauza unei omisiuni necugetate, fie din cauza proiectării intenţionate, are potenţialul de a deveni abuziv. Cămătarii care dau credite „doar cu buletinul" îl pot utiliza ca să vizeze în mod direct persoanele cărora nu le-ar fi acordate niciodată împrumuturi avantajoase fiindcă nu întrunesc condiţiile; reclamele la pariuri ar putea fi direcţionate către persoanele care frecventează site-urile cu jocuri de noroc... Şi există chiar temeri că acest gen de

* *Wish list* (engl.), listă pe care o persoană o alcătuieşte pentru a-i ajuta pe cei care ar dori să îi facă daruri să aleagă obiecte pe care chiar şi le doreşte. În anumite ţări există obiceiul ca oamenii să facă schimb de astfel de liste cu diverse ocazii: Crăciun, nuntă, botez etc. Un *gift registry* (vezi prima notă de la p. 48) este un caz special de listă cu lucruri dorite (n. red.).

elaborare a profilului cu ajutorul informațiilor poate fi folosit chiar și împotriva populației: astfel, cei pasionați de motociclete pot fi considerați a avea un hobby periculos, iar persoanele care consumă dulciuri fără zahăr pot fi semnalate ca fiind diabetice și tuturor li se poate refuza încheierea unei asigurări din aceste cauze.

Un studiu din 2015 a arătat că Google afișa mult mai puține reclame la posturi de conducere bine plătite femeilor care navigau pe internet, decât bărbaților.[12]

Iar după ce o profesoară afro-americană de la Harvard a descoperit că, atunci când își căuta numele pe Google, primea după aceea reclame care vizau persoane cu antecedente penale (și, drept urmare, a fost nevoită să-i demonstreze unui posibil angajator că nu avusese niciodată probleme cu poliția), aceasta a început să cerceteze reclamele primite de diferite grupuri etnice. Din câte a descoperit, căutările pentru „nume cu rezonanță africană" aveau șanse disproporționat de mari să fie însoțite de reclame care conțineau cuvântul „arest" (de exemplu: „Ați fost arestat?"), decât acelea pentru „numele care păreau să le aparțină unor albi".[13]

Aceste metode nu sunt folosite numai de către brokerii de date, căci există diferențe foarte mici între modul în care lucrează cei din urmă și acela în care operează Google, Facebook, Instagram și Twitter. Acești giganți ai internetului nu fac bani pentru că atrag utilizatori, așa că modelele lor de afaceri se bazează pe ideea de marketing personalizat (*micro-targeting*). Toate aceste platforme sunt, de fapt, niște motoare uriașe de furnizare de reclame, câștigând bani din faptul că milioane de utilizatori, angajați activ pe site-urile lor, dau clicuri în stânga și în dreapta, citesc postări sponsorizate, urmăresc videoclipuri sponsorizate, privesc fotografii sponsorizate.

În orice colțișor de internet pe care-l veți accesa vor sta ascunși în fundal niște algoritmi care comercializează informații despre dumneavoastră pe care nici nu știați că le deține cineva și pe care nu le-ați oferit nicicând nimănui de

bunăvoie. Acești algoritmi ne-au transformat secretele cele mai personale și mai intime într-o marfă.

Din nefericire, în numeroase țări legea nu face prea mult ca să te protejeze. Activitatea brokerilor de date este, în marea sa majoritate, nereglementată și – mai ales în America – ocaziile de a le mai reduce puterea pe care o dețin au fost ignorate în mod repetat de către guvern. În martie 2017, de exemplu, Senatul Statelor Unite a votat pentru eliminarea reglementărilor care i-ar fi împiedicat pe brokerii de date să vă vândă istoricul navigării pe internet în lipsa consimțământului dumneavoastră explicit. Aceste reglementări fuseseră, în prealabil, aprobate în octombrie 2016 de către Comisia Federală de Comunicații, însă, după schimbarea care a avut loc în guvern la sfârșitul acelui an*, au fost respinse de noua majoritate republicană a FCC (Federal Communications Commission)** și de republicani, în Congres.[14]

Însă ce înseamnă, mai exact, toate acestea pentru confidențialitatea dumneavoastră? Ei bine, dați-mi voie să vă relatez o investigație condusă de jurnalista germană Svea Eckert și analistul de date Andreas Dewes, care ar trebui să vă ofere o imagine mai clară.[15]

Eckert și echipa ei au inventat un fals broker de date și l-au folosit pentru a cumpăra datele referitoare la navigarea anonimă pe net a trei milioane de cetățeni germani. (A fost ușor să intre în posesia istoricului navigării pe internet a acestora. O mulțime de companii aveau de vânzare o groază de informații de acest gen despre clienți britanici sau americani – singura provocare a fost găsirea datelor care aveau legătură cu Germania.) Informațiile în sine fuseseră culese de un *plug-in*

* La 8 noiembrie 2016 au avut loc în SUA alegerile prezidențiale al căror câștigător a fost Donald Trump. (n. tr.).

** Comisia Federală de Comunicații, agenție independentă a guvernului american însărcinată cu reglementarea telecomunicațiilor din statele americane (n. red.).

al Google Chrome pe care utilizatorii îl descărcaseră de bunăvoie, în completă necunoștință de cauză că acesta îi mai și spiona pe parcurs.*

În total, s-a obținut o listă imensă de URL-uri. O înregistrare a tot ceea ce respectivele persoane văzuseră on-line în decursul unei luni. Fiecare căutare, fiecare pagină, fiecare clic. Totul, pus în mod legal la vânzare.

Pentru Svea Eckert și colegii ei, singura problemă a fost că informațiile de navigare erau anonime. Ceea ce ar fi trebuit să fie niște vești bune pentru toți oamenii al căror istoric de browsing fusese vândut, nu-i așa? Căci, în cazul acesta, niciunul dintre ei n-ar mai fi trebuit să se înroșească de rușine, corect? Nu, nu! Greșit! Așa cum a explicat echipa de investigații într-o prezentare la DEFCON în anul 2017, dezanonimizarea imenselor baze de date cu istoricele de navigare pe internet a fost fenomenal de simplă.

Iată cum au reușit s-o facă. Uneori, au existat indicii directe despre identitatea persoanei chiar în URL-uri. De exemplu, toți cei care au vizitat Xing.com, echivalentul german al LinkedIn. Dacă dai clic pe fotografia de profil a website-ului Xing, ești trimis la o pagină cu o adresă de genul:

www.xing.com/profile/Hannah_Fry?sc_omxb_p

Numele te trădează imediat, iar textul care urmează după numele de utilizator înseamnă că utilizatorul este logat și își vizionează propriul profil, astfel că echipa putea fi convinsă că persoana respectivă își privea pagina personală. Același lucru este valabil și în cazul lui Twitter. Oricine își verifica propria pagină de Twitter Analytics

* Acel *plug-in*, ironic denumit *The Web of Trust* („Rețeaua Încrederii"), stipula asta în mod clar, negru pe alb, ca parte a termenilor și condițiilor (n. a.).

se dădea de gol prin această acțiune. Pentru cei în cazul cărora în setul de date deținut nu exista și un identificator care să poată fi detectat instantaneu, echipa avea un alt as în mânecă... Toți cei care postaseră vreodată un link on-line (poate printr-un tweet despre un site sau prin partajarea playlist-ului de pe YouTube), practic, oricine lăsase în public o urmă digitală, o „umbră" informatică legată de numele său adevărat, a fost demascat în mod accidental prin această acțiune. Echipa de cercetători a folosit un algoritm simplu pentru a pune în corespondență persoanele reale cu cele anonime[16], filtrându-le lista de URL-uri pentru a găsi pe cineva din setul de date care vizitase aceleași site-uri la aceleași date calendaristice și la aceleași repere orare la care fuseseră postate linkurile on-line. În cele din urmă, au obținut numele și prenumele tuturor persoanelor din setul de date, precum și accesul complet la istoricul de navigare dintr-o lună de zile a milioane de germani.

Printre aceste trei milioane de oameni se numărau și câteva persoane importante: un politician care căutase niște medicamente on-line. Un ofițer de poliție care copiase un document dintr-un dosar delicat și-l lipise într-o fereastră din Google Translate, iar detaliile acestuia apăruseră apoi în URL, putând fi văzute de cercetători. Și un judecător al cărui istoric de căutare dezvăluia vizitarea zilnică a unui domeniu destul de „specializat" de pe internet. Iată, în cele ce urmează, o mică selecție a site-urilor pe care le vizitase acesta într-un interval de opt minute, în august 2016:

18.22: http://www.tubegalore.com/video/amature-pov-ex-wife-in-leather-pants-gets-creampie42945.html

18.23: http://www.xxkingtube.com/video/pov_wifey_on_sex_ stool_with_beaded_thong_gets_creampie_4814.html

18.24: http://de.xhamster.com/movies/924590/office_lady_in_ pants_rubbing_riding_best_of_anlife.html

18.27: http://www.tubegalore.com/young_tube/5762–1/page0

18.30: http://www.keezmovies.com/video/sexy-dominatrix-milks-him-dry-1007114?utm_sources

Pe lângă aceste sesiuni zilnice de navigare pe internet, judecătorul mai căuta, cu regularitate, on-line nume de nou-născuți, cărucioare și spitale cu maternitate. Echipa a tras concluzia că în perioada aceea partenera sa aștepta un copil.

Să lămurim, însă, un lucru: judecătorul nu făcea nimic ilegal. Mulți oameni – printre care mă număr și eu – vor susține că nu făcea nimic rău. Însă nu este mai puțin adevărat că aceste informații ar fi fost utile pentru cineva care ar fi vrut să-l șantajeze sau să-i pună familia într-o situație stânjenitoare.

Iar acesta este locul de unde începem să ne îndepărtăm și să depășim granița care ne desparte de lucrurile sinistre – atunci când informații private și delicate despre propria persoană, culese fără știrea ta, sunt folosite ulterior pentru a te manipula. Ceea ce, evident, este exact ce a făcut firma britanică de consultanță politică Cambridge Analytica.

CAMBRIDGE ANALYTICA

Probabil că deja cunoașteți cea mai mare parte a poveștii.

Începând din anii 1980, psihologii folosesc un sistem de cinci caracteristici pentru cuantificarea personalității unui individ. Primești un punctaj pentru fiecare dintre următoarele trăsături: deschiderea (*openness*) față de noile experiențe, conștiinciozitate (*conscientiousness*), caracter extravertit (*extraversion*), caracter agreabil (*agreeableness*) și caracter

nevrotic (*neuroticism*).* Luate împreună, acestea oferă o metodă standard deosebit de utilă pentru a descrie ce tip de persoană ești.

În 2012, cu un an înainte de intrarea în scenă a Cambridge Analytica, mai mulți oameni de știință de la Universitatea Cambridge și de la Universitatea Stanford au început să caute o legătură între cele cinci trăsături de personalitate și paginile cărora oamenii le dădeau like-uri pe Facebook.[17] Au construit un chestionar pe Facebook în acest scop, lăsând utilizatorii să efectueze teste reale de psihometrie în speranța că vor descoperi o legătură între caracterul autentic al unei persoane și identitatea sa on-line. Cei care au descărcat chestionarul le-au oferit, cu bună știință, informații și despre istoricul like-urilor personale de pe Facebook, iar, prin intermediul răspunsurilor la o serie de întrebări, valorile personalității lor reale, procentele în care dețineau cele cinci caracteristici.

Este ușor să ne imaginăm cum ar putea fi corelate like-urile date cu personalitatea. Așa cum a subliniat echipa de cercetători în lucrarea pe care a publicat-o în anul următor[18], aceia cărora le plac Salvador Dalí, meditația sau conferințele de pe platforma TED vor obține aproape sigur punctaje mari la deschiderea către noi experiențe. În schimb, persoanele cărora le plac petrecerile, dansul și personajul Snooki din serialul de televiziune *Jersey Shore* tind să fie ceva mai extravertite. Cercetarea a fost încununată de succes. După

* Cunoscut în psihologie drept modelul celor cinci trăsături sau modelul celor cinci factori (*five-factor model*, FFM, engl.) ori modelul OCEAN – acronim al denumirilor în engleză ale celor cinci trăsături ale personalității: *openness* (engl., „deschidere"), *conscientiousness* (engl., „conștiinciozitate"), *extraversion* (engl., „caracter extrovertit"), *agreeableness* (engl., „caracter agreabil) și *neuroticism* (engl., „caracter nevrotic") –, reprezintă un model de descriere a spectrului personalităților propus în 1980 de psihologii Lewis Goldberg, Naomi Takemoto-Chock, Andrew Comrey și John M. Digman și rafinat de atunci de mai mulți cercetători (n. red.).

ce a stabilit această legătură, echipa a construit un algoritm care putea deduce personalitatea cuiva numai din like-urile pe care le dădea pe Facebook.

Până să le apară cel de-al doilea studiu, în 2014[19], cercetătorii susțineau deja că, dacă reușeai să strângi 300 de like-uri din profilul de pe Facebook al unei persoane, algoritmul putea să-i analizeze acesteia caracterul mai precis decât soțul sau soția sa.

Revenind în prezent, trebuie spus că acest grup dedicat cercetării științifice, mai precis, Centrul de Psihometrie al Universității Cambridge, a extins algoritmul pentru a face predicții despre personalitatea cuiva și cu ajutorul mesajelor postate de acesta pe Twitter. Există un site, accesibil oricui, unde poți să încerci chiar tu algoritmul. Din moment ce profilul meu pe Twitter este public, m-am gândit să încerc predicțiile cercetătorilor asupra propriei persoane, așa că le-am trimis istoricul mesajelor de pe Twitter și, pentru a avea un termen de comparație, am completat un studiu de personalitate bazat pe un chestionar tradițional. Algoritmul a reușit să-mi evalueze corect trei însușiri din cinci, cu toate că, din câte s-a dovedit din studiul tradițional de personalitate, sunt mult mai extravertită și nevrotică decât lasă profilul meu de pe Twitter să se înțeleagă.*

Toate aceste activități au fost motivate de faptul că ele puteau fi de folos în publicitate. Astfel, în 2017[20], aceeași echipă de cadre universitare a continuat cu un alt experiment, trimițând reclame adaptate la trăsăturile de personalitate ale unei persoane. Folosind platforma Facebook, echipa a furnizat reclame la produse de înfrumusețare persoanelor extravertite, folosind sloganul: „Dansează ca și când nu te-ar vedea nimeni (deși te vede toată lumea)", pe când persoanele introvertite au văzut imaginea unei fete care își zâmbea, privindu-se în

* Această combinație pare să sugereze că aș posta mai multe lucruri dacă nu mi-aș face atâtea griji în legătură cu felul în care vor fi primite de ceilalți (n. a.).

oglindă, însoțită de textul: „Frumusețea nu trebuie neapărat să fie stridentă."

Într-un experiment desfășurat în paralel, persoanelor din grupul-țintă care erau într-o mare măsură deschise spre acceptarea experiențelor noi li s-au arătat reclame la cuvinte încrucișate folosind o imagine cu textul: „Aristotel? Insulele Seychelles? Dați-vă frâu liber creativității și stimulați-vă imaginația cu o cantitate nelimitată de cuvinte încrucișate!" Aceleași cuvinte încrucișate erau apoi prezentate pe ecran persoanelor care erau doar într-o măsură redusă deschise spre experiențe noi, dar folosindu-se îndemnul: „Instalați-vă comod alături de jocul preferat dintotdeauna! Cuvintele încrucișate care au constituit o provocare pentru generații întregi de pasionați." Per ansamblu, echipa a susținut că stabilirea unei corelații între reclame și caracterul unei persoane a generat cu până la 40% mai multe clicuri și cu până la 50% mai multe achiziții decât folosirea reclamelor generice, nepersonalizate. Iar pentru orice agenție de publicitate, acestea sunt rezultate de-a dreptul impresionante!

În tot acest timp, pe când cadrele universitare își publicau lucrarea, alții le implementau metodele. Printre ei s-a numărat și compania Cambridge Analytica în perioada în care a lucrat pentru campania electorală a lui Donald Trump.

Haideți, acum, să ne întoarcem puțin în timp! Există puține dubii că, de fapt, Cambridge Analytica ar fi folosit alte strategii decât fictiva mea agenție de voiaje de lux, Fry's. Abordarea lor a constat în identificarea unor grupuri reduse de persoane pe care le considerau ușor de convins și cărora le trimiteau apoi reclame țintite, vizându-le direct, în loc să le trimită reclame nediferențiate, pe baza unui contract-cadru*. S-a descoperit, de exemplu, că anumite categorii se suprapuneau într-o mare măsură, spre exemplu, aceea a persoanelor care cumpăraseră

* *Blanket advertising* (engl.), un aranjament de cumpărare de reclame, în care o agenție de publicitate se obligă să distribuie un anumit număr, la un anumit preț fix, pe o perioadă determinată. (n. tr.).

produse americane get-beget precum autoturisme Ford şi aceea a cetăţenilor care se înregistraseră în registrele de votare ca susţinători ai Partidului Republican. Aşadar, cei de la Cambridge Analytica au început să caute persoane care preferau marca Ford, dar despre care nu se ştia şi că ar fi votat cu republicanii, să vadă dacă puteau să le influenţeze opiniile prin utilizarea de reclame naţionaliste, care să le stârnească patriotismul. Într-un fel, asta nu diferă prea mult de strategia folosită de un candidat care identifică un cartier cu alegători indecişi şi se duce din uşă în uşă pentru a-i convinge pe rând. Iar în materie de campanii on-line, nu se deosebeşte de ceea ce au făcut Obama şi Clinton în timpul propriilor campanii electorale. Fiecare partid politic important din lumea occidentală foloseşte analize extensive ale datelor informatice şi diverse modalităţi de micro-targetare a alegătorilor.

Dacă este să dăm crezare reportajului sub acoperire înregistrat de *Channel Four News*, atunci Cambridge Analytica a folosit şi strategia realizării de profiluri de personalitate ale electoratului pentru difuzarea de mesaje politice cu mare încărcătură emoţională – de exemplu, au descoperit mame singure care aveau într-un procent mai mare un caracter nevrotic şi au profitat de teama lor de a fi atacate în propria locuinţă, pentru a le convinge să susţină un mesaj politic care făcea lobby pentru deţinerea armelor de foc. Agenţiile de publicitate au folosit în mod cert pe scară largă aceste strategeme şi probabil că la fel au făcut-o şi coordonatorii altor campanii politice.

Numai că pe lângă asta, Cambridge Analytica este acuzată şi de crearea de reclame şi mascarea lor în jurnalism. Conform declaraţiilor unui denunţător, preluate de ziarul britanic *Guardian*, una dintre cele mai eficiente reclame din timpul campaniei a fost un grafic interactiv intitulat: „10 adevăruri incomode despre Fundaţia Clinton".[21] Un alt denunţător a mers şi mai departe, susţinând că „articolele"

plantate de Cambridge Analytica în presă se bazau, adesea, pe minciuni demonstrabile.[22]

Să presupunem, de dragul discuției, că toate cele de mai sus ar fi adevărate: Cambridge Analytica a transmis diverse știri false (*fake news*) manipulatoare pe Facebook anumitor persoane, în funcție de profilurile psihologice ale acestora. Întrebarea este: a dat roade strategia lor?

MICRO-MANIPULAREA

Există o totală lipsă de simetrie în modul în care ne raportăm la puterea pe care o dețin reclamele politice țintite. Ne place să credem că noi înșine avem o gândire independentă, că suntem imuni la manipulare și, totuși, ni-i imaginăm pe „ceilalți" – în special, pe cei care au convingeri politice diferite – ca fiind inimaginabil de creduli. Adevărul este, probabil, undeva, la mijloc.

Ceea ce știm în mod sigur este că postările pe care le vedem pe Facebook au capacitatea de a ne modifica emoțiile. În cadrul unui experiment controversat, realizat de angajați ai Facebook în anul 2013, a manipulat fluxurile de știri care li se afișau unui număr de 689 003 utilizatori fără cunoștința (sau consimțământul) acestora, în încercarea de a le controla emoțiile și de a le influența starea de spirit.[23] Cei care au realizat cercetarea au blocat toate postările prietenilor respectivilor utilizatori care conțineau cuvinte cu semnificații pozitive, după care au făcut același lucru cu postările care includeau cuvinte negative și au urmărit cum reacționau subiecții neavizați în fiecare dintre aceste situații. Utilizatorii care au observat mai puțin conținut negativ în fluxurile lor au început să posteze ei înșiși lucruri mai degrabă pozitive. Între timp, cei cărora le fuseseră ascunse din *timeline* postările pozitive au început să folosească ei înșiși cuvinte negative. Concluzia: ne

putem imagina c-am fi imuni la manipularea emoțională, dar probabil că nu suntem.

Mai știm, din experimentul lui Epstein descris în capitolul „Puterea", că până și ordinea paginilor care afișează rezultatele unui motor de căutare poate fi suficientă pentru a-i influența pe alegătorii indecișI să devină favorabili unui anumit candidat și nu altuia. Cunoaștem, de asemenea, din activitatea acelorași cercetători universitari ai căror algoritmi Cambridge Analytica i-a utilizat într-un scop diferit, că reclamele sunt mai eficiente atunci când vizează trăsăturile de personalitate ale indivizilor.

Toate aceste informații însumate constituie un argument puternic pentru a sugera că astfel de metode pot influența modul în care oamenii votează, tot așa cum au o influență asupra modului în care își cheltuiesc banii. Dar – și este un mare „dar" –, mai trebuie să știți ceva înainte de a trage o concluzie.

Toate cele de mai sus sunt adevărate, numai că efectele propriu-zise sunt nesemnificative. În experimentul realizat de cei de la Facebook, utilizatorii au fost, într-adevăr, mai înclinați să posteze mesaje pozitive dacă li se ascundeau știrile cu conținut negativ. Numai că diferența era de mai puțin de o zecime din 1%!

În mod asemănător, în exemplul cu reclamele personalizate, produsele cosmetice vândute introverților aveau mai mult succes dacă se luau în considerare trăsăturile de personalitate ale unei persoane, însă diferența obținută era infimă. O reclamă generică determina 31 de persoane dintr-o mie să dea clic pe ea. Chiar și procentul acela de 50% îmbunătățire pe care l-am citat mai devreme, care era îndrăzneț întipărit la începutul lucrării științifice, se referă, de fapt, la o creștere infimă, de la 11 clicuri la 1 000 de persoane care văzuseră reclama, la 16!

Strategiile pot să dea rezultate, este adevărat. Dar companiile de publicitate nu-și injectează mesajele direct în mințile unei audiențe pasive. Nu suntem niște ținte chiar atât de ușoare. Ne pricepem mult mai bine să ignorăm reclamele

sau să interpretăm în felul nostru mesajele propagandiste decât și-ar dori cei care ni le trimit. Până la urmă, chiar și în cazul celor mai bune și mai parșive campanii publicitare axate pe micro-profilare, grupul-țintă va ajunge să fie influențat într-o măsură foarte mică.

Cu toate acestea, la nivel teoretic, cel puțin, într-o campanie electorală, astfel de frânturi de capacitate de influențare ar putea fi suficiente pentru a înclina balanța în favoarea unui candidat sau a altuia. La o populație de zeci sau sute de milioane de indivizi, acele schimbări de opinie de unu la mie se pot însuma cu repeziciune. Și, când îți amintești că, așa cum a subliniat Jamie Bartlett într-un articol pentru revista britanică *Spectator*, Trump a câștigat alegerile din statul Pennsylvania cu o diferență de 44 000 de voturi „pentru" din 6 000 000 exprimate, pe cele din Wisconsin, cu 22 000 de voturi, iar pe cele din Michigan cu doar 11 000 de voturi, atunci s-ar putea ca până și niște marje de influențare de sub 1% să fie exact ceea ce-ți trebuie.[24]

Adevărul este că nu se poate preciza ce efect au avut toate acestea în alegerile prezidențiale din Statele Unite. Chiar dacă am avea acces la toate faptele, nu ne putem întoarce în timp, să descâlcim păienjenișul de cauze și efecte pentru a indica fie și un singur motiv pentru decizia de a vota a vreunei persoane. Ce-a fost a fost și nu mai poate fi schimbat. Ce contează este spre ce ne îndreptăm în viitor.

EVALUEAZĂ-MĂ!

Este important să ne amintim că am beneficiat cu toții de acest model pe care este construit internetul. Peste tot în lume, oamenii au acces liber și gratuit la rețelele globale de comunicare instantanee, având tot tezaurul cunoașterii umane la îndemână, informații de ultimă oră din toate colțurile pământului și utilizare nelimitată a celor mai

remarcabile programe de calculator și a celor mai avansate tehnologii, realizate de companii private – totul plătit de publicitate. Aceasta este înțelegerea pe care am încheiat-o. Tehnologie gratuită, în schimbul posibilității ca aceasta să fie folosită pentru a ne influența și pentru a se profita de pe urma noastră. Tot ce este mai bun și tot ce este mai rău din capitalism dintr-o singură mișcare simplă.

Am putea să ajungem la concluzia că suntem fericiți cu târgul pe care l-am încheiat. Și asta este cât se poate de în regulă. Dacă îl acceptăm, însă, este important să fim conștienți de la bun început de toate pericolele pe care le poate genera strângerea atâtor informații. Trebuie să avem în minte la ce ar putea duce aceste baze de date în viitor– chiar dincolo de problemele legate de protecția datelor și de potențialul lor de a submina democrația (de parcă acestea nu ar fi, oricum, îndeajuns de rele!). Dar există și o altă răsturnare de situație în această poveste distopică: o aplicație creată pentru a folosi astfel de baze de date detaliate, uriașe și intercorelate, care pare demnă să fie prezentată într-un episod al popularului serial SF de pe Netflix *Black Mirror**, dar care există în realitate. Este cunoscută drept Sesame Credit, este utilizată de guvernul chinez și reprezintă un sistem care acordă un punctaj fiecărui cetățean.

Imaginați-vă că fiecare informație pe care un broker de date ar avea-o despre dumneavoastră ar fi redusă la un punctaj unic. Absolut tot a fi inclus: cât de responsabil (sau de irresponsabil) v-ați achitat creditele, numărul telefonului mobil, adresa – lucrurile obișnuite. Dar și comportamentul de zi cu zi, postările de pe rețelele de socializare, informațiile adunate de aplicația prin care solicitați taxiuri, chiar și înregistrări ale

* „Nosedive", primul episod al celui de-al treilea sezon al serialului *Black Mirror* (în regia lui Joe Wright și scenariul lui Michael Schur și Rashida Jones, după o poveste de Charlie Brooker), a fost adesea considerat o metaforă pentru Social Credit System, sistemul de credit social implementat de guvernul chinez, preuspoziție confirmată de producătorul Charlie Brooker într-o conferință de presă (n. red.).

activității on-line pe un site de anunțuri matrimoniale. Rezultatul este un singur număr cuprins între 350 și 950 de puncte.

Sesame Credit nu dezvăluie detalii despre algoritmul său „complex" de punctare. Însă Li Yingyun, directorul serviciului de tehnologie, a oferit câteva exemple despre ceea ce s-ar putea deduce din rezultatele obținute, într-un interviu luat de *Caixin Media*, grupul mediatic cu sediul în Beijing:

> Cineva care joacă jocuri video zece ore pe zi, de exemplu, va fi considerat o persoană inactivă. Cineva care cumpără frecvent scutece pentru copii va putea fi considerat părinte, care, prin comparație, este mai probabil să posede un anumit simț al responsabilității.[25]

Iar pentru chinezi, acest scor chiar a ajuns să conteze. Dacă punctajul tău depășește 600 de puncte, poți să obții un card de credit special. Dacă depășește 666 de puncte, ești recompensat cu o limită de credit superioară. Cei cu punctaje peste 650 pot să închirieze o mașină fără să depună garanție și să folosească un culoar pentru VIP-uri la aeroportul din Beijing. Oricine are un punctaj de peste 750 poate să solicite în regim de urgență o viză de intrare în Europa.[26]

Totul este distractiv acum, cât timp sistemul de acordare a unor punctaje cetățenilor este încă opțional. Dar în 2020, când va deveni obligatoriu, persoanele cu punctaje mici vor începe să simtă repercusiunile în fiecare aspect al vieții lor. Chiar documentele guvernului cu privire la acest sistem prezintă exemple de sancțiuni, care s-ar putea aplica oricărei persoane considerate recalcitrante:

> Restricții în privința plecării în străinătate, restricții de cumpărare a unor [...] bunuri, de a călători cu avionul, în scop de turism și în vacanțe, sau restricții de cazare în hoteluri cu un anumit număr de stele.

Documentele oficiale mai avertizează că, în cazul unor „subiecți care vor înșela grav încrederea [partidului]", acesta va „sfătui băncile comerciale [...] să le limiteze respectivilor

accesul la credite, la cumpărarea de asigurări şi la alte servicii financiare de acelaşi gen".[27] Loialitatea este elogiată. Înşelarea încrederii este pedepsită. Aşa cum se exprima Rogier Creemers, un cadru universitar specializat în dreptul şi guvernarea chineză care predă la Institutul „Van Vollenhoven" al Universităţii Leiden:

> Modalitatea cea mai bună de a înţelege acest sistem este să-l considerăm un soi de bastard, rod al iubirii pătimaşe, al unui sistem de fidelizare a clienţilor.[28]

Nu pot să vă ofer informaţii prea liniştitoare în privinţa lui Sesame Credit, dar nu doresc nici să credeţi că perspectivele noastre sunt cu totul şi cu totul sumbre, căci se întrezăreşte o rază de speranţă în altă parte. Oricât de mohorât ar părea viitorul, sunt semnale că, încetul cu încetul, direcţia vântului se schimbă. Mulţi dintre cei care fac parte din comunitatea specialiştilor în ştiinţa datelor sunt, de ceva timp, la curent cu şi obiectează la exploatarea informaţiilor despre oameni în schimbul obţinerii de profituri. Până la scandalul stârnit de Cambridge Analytica, însă, aceste chestiuni nu au atras atenţia în mod special, neajungând, nicăieri în lume, pe prima pagină a ziarelor. La începutul lui 2018 însă, când a izbucnit scandalul, publicul larg a văzut pentru întâia oară cum algoritmii culeg, pe tăcute, informaţii despre oameni şi şi-a dat seama că, în lipsa unei supravegheri atente sau a unor reglementări, acest lucru ar putea avea repercusiuni dramatice.

Iar reglementările nu întârzie să sosească. Dacă locuiţi în Uniunea Europeană, aici a fost adoptată de curând o nouă lege, numită GDPR (General Data Protection Regulation*), care decretează a fi ilegale destul de multe dintre activităţile

* Regulamentul general privind protecţia datelor (denumire completă: Regulamentul nr. 679 din 27 aprilie 2016 privind protecţia persoanelor fizice în ceea ce priveşte prelucrarea datelor cu caracter personal şi libera circulaţie a acestor informaţii şi de abrogare a Directivei 95/46/CE) (n. tr.).

întreprinse de către brokerii de date. Teoretic, acestora nu li se va mai permite să vă stocheze informațiile fără un scop explicit. Nu vor putea să deducă informații despre persoana dumneavoastră în lipsa consimțământului exprimat explicit. Și nu vor mai putea să vă ceară permisiunea de a vă colecta datele personale dintr-un anumit motiv, iar apoi să le folosească în secret pentru altceva. Asta nu înseamnă neapărat că acest gen de practici vor și înceta. În primul rând, pentru că, de cele mai multe ori, noi nici nu dăm atenție termenilor și condițiilor pentru care semnăm că suntem de acord atunci când facem clic on-line în stânga și în dreapta, așa că ne-am putea găsi în situația de a ne exprima consimțământul chiar fără să ne dăm seama. În al doilea rând, identificarea practicilor ilegale și aplicarea reglementărilor rămâne o chestiune complicată într-o lume în care cea mai mare parte a analizei și transferului de date se face din umbră. Prin urmare, trebuie să așteptăm și să vedem cum vor decurge lucrurile.

Europenii sunt mai norocoși, dar și în Statele Unite există persoane care fac presiuni pentru reglementarea situației. În 2014, Federal Trade Comission (Comisia Federală pentru Comerț) a publicat un raport în care condamna practicile dubioase ale brokerilor de date, iar de atunci a exercitat tot felul de presiuni pentru a spori drepturile consumatorilor. Apple a integrat deja un soft pentru „protecție inteligentă împotriva urmăririi"* în browser-ul său, Safari, iar Firefox a făcut același lucru. Facebook își întrerupe legăturile cu brokerii săi de informații. Argentina și Brazilia, Coreea de Sud și multe alte țări au impus legi asemănătoare GDPR-ului. Poate că Europa este cu un pas înainte, dar se manifestă o tendință globală de a încerca să ne îndreptăm în direcția cea bună.

Dacă datele personale informatizate sunt, cu adevărat, noul aur înseamnă că, până acum, am trăit în Vestul Sălbatic.

* Intelligent Tracking Prevention sau ITP (n. tr.).

Eu, una, sunt, însă, optimistă și cred că, pentru cei mai mulți dintre noi, ce-a fost mai rău a trecut deja.

Am face, totuși, bine să ne amintim că în viață nimic nu este gratis. În timp ce legislația încearcă să recupereze decalajul față de realitatea tehnologică și se poartă, ca întotdeauna, bătălia cu dorința de profit a corporațiilor pentru obținerea a ceeea ce este bine pentru societate, trebuie să fim atenți să nu ne amăgim cu falsa impresie că am avea parte de confidențialitate. Ori de câte ori folosim un algoritm – mai ales unul care ne este oferit fără să plătim! – trebuie să ne întrebăm care sunt motivațiile ascunse ce stau la baza acestei acțiuni. De ce această aplicație îmi oferă astfel de chestii în mod gratuit? Ce face, de fapt, algoritmul acesta? Este, oare, un schimb care nu mă deranjează? Mi-ar fi mai bine fără el?

Este o lecție valabilă și dincolo de mediul virtual, pentru, că, în prezent, influența acestui gen de calcule se extinde, practic, asupra oricărui aspect al societății. Informațiile și algoritmii nu mai dețin doar puterea de a prezice obiceiurile la shopping ale cuiva. Ele pot, de asemenea, să priveze o persoană de liberate.

Capitolul 3

JUSTIȚIA

Vara, nu este deloc neobișnuit să dai peste petrecăreți inofensivi care beau câte un pahar, într-o seară de duminică, pe străzile din Brixton, acolo unde începe următoarea noastră poveste. Acesta este un district din sudul Londrei, care are reputația de a fi un loc bun în care să te distrezi cu prietenii. Iar într-una din seri tocmai se încheiase un festival de muzică, iar zona era plină de oameni care se pregăteau, binedispuși, să se întoarcă acasă sau să continue cheful. Pe la 23.30, însă, starea lor de spirit s-a schimbat. În curtea unei locuințe sociale a izbucnit o bătaie, iar când polițiștii nu au reușit să țină situația sub control, totul s-a extins rapid până în centrul localității, unde sute de tineri li s-au alăturat recalcitranților.

Era în august 2011. Cu o noapte înainte, în cealaltă parte a orașului, un protest care începuse pașnic împotriva împușcării de către poliție a unui tânăr din Tottenham pe nume Mark Duggan se transformase într-o confruntare violentă. Acum, pentru a doua noapte la rând, zone întregi din oraș erau cuprinse de haos, iar de data aceasta atmosfera era diferită. Ceea ce începuse ca o demonstrație locală se transformase repede într-o amplă manifestare a sfidării legii și ordinii și o devalizare generalizată.

Chiar când răzmerița începea să ia amploare, Nicholas Robinson, un student la Electrotehnică în vârstă de 23 de ani ce își petrecuse sfârșitul de săptămână la prietena lui, se întorcea

acasă folosind scurtătura sa obişnuită prin Brixton.[1] De acum, străzile familiare deveniseră, practic, de nerecunoscut: de o parte şi de alta a drumului se zăreau maşini răsturnate, ferestre sparte, case incendiate şi magazine vandalizate.[2] Poliţiştii se străduiau, disperaţi, să calmeze situaţia, dar le era cu neputinţă să oprească maşinile şi scuterele care veneau chiar lângă vitrinele sparte şi încărcau bunurile de furat: haine, încălţăminte şi televizoare. Brixtonul scăpase complet de sub control.

La câteva străzi distanţă de un magazin de produse electronice care fusese complet devastat, Nicholas Robinson a trecut pe lângă supermarketul unde îşi făcea de obicei cumpărăturile. Aproape la fel ca şi celelalte magazine, şi acesta era un dezastru: vitrinele şi uşile din sticlă fuseseră sparte, iar rafturile din interior erau vraişte din cauza distrugerilor provocate de jefuitori. Valuri de insurgenţi treceau în fugă pe lângă el, ţinând strâns în braţe laptopuri noi-nouţe, fără a fi opriţi de poliţie. În mijlocul haosului, simţind că-i este sete, Nicholas a intrat în magazin şi a luat un bax cu sticle de apă în valoare de 3,50 lire sterline. Şi, tocmai când dădea colţul, pregătindu-se să plece, poliţiştii au dat buzna în supermarket. Nicholas şi-a dat imediat seama ce făcuse, a lăsat să-i cadă baxul cu sticle şi a încercat s-o zbughească de-acolo.[3]

Când a sosit noaptea de luni, ţara se pregătea să înfrunte noi răzmeriţe. Şi, într-adevăr, în noaptea aceea, jefuitorii au ieşit din nou pe străzi.[4] Printre aceştia se număra şi tânărul de 19 ani Richard Johnson. Intrigat de ceea ce văzuse la ştiri, pusese mâna pe o cagulă (un articol cu totul nepotrivit pentru sezonul de vară!), sărise în maşină şi-şi croise drum până la centrul comercial din apropiere. Cu faţa ascunsă, Richard a pătruns în fugă în magazinul de jocuri pe computer al oraşului, a adunat un braţ de jocuri pe calculator şi s-a întors la maşină.[5] Din păcate pentru el, parcase chiar în faţa unei camere CCTV*. Numărul de pe plăcuţa de înmatriculare a

* *Closed-circuit television* (engl.), „camere de televiziune cu circuit închis", folosite pentru supravegherea video a unei zone (n. tr.).

mașinii a ajutat poliția să-l urmărească, iar înregistrarea video a făcut mai ușoară deschiderea unui proces împotriva lui.[6]

Atât Richard Johnson, cât și Nicholas Robinson au fost arestați pentru faptele comise de ei în timpul revoltelor din 2011. Ambii au fost acuzați de tâlhărie. Ambii s-au prezentat în fața judecătorilor. Ambii s-au declarat vinovați. Însă aici iau sfârșit asemănările dintre cele două procese.

Nicholas Robinson a fost chemat cel dintâi pe banca acuzaților, apărând în fața unui judecător de la Curtea Magistraților din Camberwell la mai puțin de o săptămână de la incident. În pofida valorii mici a sticlelor de apă pe care le furase, în pofida absenței oricăror antecedente penale, a faptului că urma cursurile învățământului la zi și că a declarat curții că îi era rușine de faptele pe care le comisese, judecătorul a decis că acțiunile sale contribuiseră la atmosfera de anarhie din Brixton din acea noapte. Și, astfel, în plânsetele membrilor familiei sale din spațiul sălii de judecată rezervat publicului, Nicholas Robinson a fost condamnat la șase luni de închisoare.[7]

Cazul lui Johnson a ajuns în instanță în ianuarie 2012. Cu toate că ieșise din casă purtând un obiect de îmbrăcăminte menit să-i ascundă identitatea din dorința deliberată de a jefui, cu toate că și el jucase un rol în agravarea dezordinii publice, Johnson a scăpat de închisoare. I s-a dat o sentință cu suspendare și a fost condamnat la efectuarea a 200 de ore de muncă neplătită în folosul comunității.[8]

DILEMA MORALĂ A CONSECVENȚEI

Sistemul judiciar știe că nu este perfect, însă nici nu încearcă să fie astfel. Aprecierea vinovăției și stabilirea pedepselor nu este o știință exactă și este imposibil ca un judecător să garanteze că sentințele sale vor fi date cu precizie. De aceea

și există expresii precum „suspiciune rezonabilă" și „motive temeinice", atât de vitale pentru vocabularul legislativ, și tot din aceste motive recursurile sunt o parte atât de importantă a proceselor: sistemul recunoaște astfel că este imposibil de atins certitudinea absolută asupra vinovăției sau nevinovăției cuiva.

Chiar și așa, discrepanțele în tratamentul aplicat unor inculpați – precum Nicholas Robinson și Richard Johnson – par incorecte. Sunt prea mulți factori implicați ca să se poată afirma cu siguranță că existența unei diferențe între sentințe este „nedreaptă", dar, în limite rezonabile, s-ar putea spera că judecătorii au fost, în general, consecvenți în modul în care au luat deciziile. Dacă, de exemplu, dumneavoastră și geamănul dumneavoastră ați comis aceeași infracțiune, ați putea spera că o instanță de judecată vă va da aceeași sentință. Însă, așa să fie, oare?

În anii 1970, un grup de cercetători americani a încercat să răspundă la o variantă a acestei întrebări.[9] În loc să folosească infractori gemeni (ceea ce este, practic, dificil și, din punct de vedere moral, indezirabil), oamenii de știință au inventat o serie de cazuri ipotetice și au întrebat, separat, 47de judecători din cadrul instanțelor districtuale ale statului Virginia cum ar proceda în fiecare dintre ele. Iată un exemplu extras din studiu, ca să încercați același lucru. Cum ați proceda în cazul următor?

> O inculpată, o tânără de 18 ani, a fost reținută pentru posesie de marijuana, fiind arestată împreună cu prietenul ei și cu alte șapte cunoștințe. Au fost găsite cantități însemnate de marijuana, fumată și nefumată, însă asupra fetei nu s-a descoperit nici urmă. Nu avea antecedente penale, era o studentă silitoare provenind dintr-o familie din clasa mijlocie și nu fusese nici sfidătoare, dar nici nu-și ceruse scuze pentru faptele ei.

Deosebirile între sentințe au fost impresionante. Din cei 47 de judecători, 29 au declarat inculpata nevinovată, iar 18 – vinovată. Dintre judecătorii care au optat pentru verdictul „vinovată", opt au recomandat eliberarea condiționată, patru

s-au gândit că o amendă era mai potrivită, trei ar fi optat și pentru eliberare condiționată, și pentru amendă, iar trei ar fi fost înclinați s-o trimită pe inculpată la închisoare.

Așadar, în baza acelorași dovezi și în cazuri identice, un inculpat se poate aștepta să scape basma curată sau să fie trimis direct la închisoare, asta depinzând cu totul de judecătorul în fața căruia a avut șansa (sau neșansa) să se prezinte.

Este o adevărată lovitură pentru oricine se mai agață încă de ideea unei justiții consecvente. Dar lucrurile stau chiar mai rău, pentru că judecătorii nu numai că sunt în dezacord unul cu altul, ci mai au și tendința să-și contrazică până și *propriile sentințe*.

Într-un studiu recent, 81 de judecători din Marea Britanie au fost întrebați dacă ar oferi eliberarea pe cauțiune unui număr de inculpați imaginari.[10] Fiecare caz ipotetic avea propria poveste fictivă și antecedente penale inventate. Asemenea confraților lor din studiul realizat în statul Virginia, judecătorii britanici nu au reușit să cadă de acord nici măcar asupra unui singur caz din cele 41 care le-au fost prezentate.[11] De data aceasta, printre cele 41 de cazuri ipotetice prezentate fiecărui judecător, șapte dintre ele apăreau de două ori – cu numele inculpaților schimbate la a doua apariție pentru ca judecătorul să nu-și dea seama că erau duplicate. Era un șiretlic parșiv, însă revelator. Cei mai mulți judecători nu au reușit să ia aceeași decizie în același dosar atunci când l-au văzut pentru a doua oară. Uimitor este că unii judecători nu s-au descurcat mai bine în a da același răspuns într-o situație identică decât dacă ar fi stabilit, efectiv, cauțiunea la întâmplare.[12]

Multe alte studii au ajuns la o concluzie identică: ori de câte ori judecătorii au libertatea de a evalua cazurile ei înșiși, vor apărea neconcordanțe majore. A lăsa judecătorilor libertate de interpretare înseamnă a permite existența elementului șansă în sistem.

Există o soluție simplă, desigur. O cale ușoară pentru a garanta consecvența judecătorilor este aceea de a le lua dreptul de a exercita puteri discreționare. Dacă fiecare persoană

acuzată de același delict ar fi tratată în exact aceeași manieră, atunci s-ar putea garanta acuratețea actului de justiție – cel puțin, în ceea ce privește stabilirea cauțiunii și a pedepsei. Și, într-adevăr, unele țări au pornit pe calea aceasta. În Statele Unite, la nivel federal, și în unele zone din Australia operează deja sisteme prescriptive de stabilirea sentințelor.[13] Însă acest gen de consecvență are, totuși, neajunsuri, căci, pentru tot ce se câștigă în precizie, se pierde în ceea ce privește un alt fel de echitate.

Pentru a ilustra ceea ce am spus, închipuiți-vă doi inculpați, ambii acuzați de furt dintr-un supermarket. Unul este un infractor de carieră, oarecum în elementul lui, care fură de bunăvoie; celălalt și-a pierdut de curând locul de muncă și se luptă să o scoată la capăt, furând ca să-și hrănească familia și fiind plin de remușcări din cauza faptelor sale. Prin eliminarea libertății de a ține cont de aceste circumstanțe atenuante, regulile stricte care îi tratează în același fel pe toți cei care se fac vinovați de aceleași infracțiuni pot sfârși prin a fi inutil de aspre pentru unii și pot însemna că se ratează orice șansă de a-i reabilita pe anumiți infractori.

Iată o adevărată dilemă morală. Oricum s-ar stabili sistemul în care să acționeze judecătorii, trebuie găsit un echilibru delicat între oferirea unei justiții individualizate și asigurarea consecvenței sentințelor. Cele mai multe țări au încercat să rezolve dilema optând pentru adoptarea unor sisteme care se situează undeva între extrema prescriptivă a legii federale din Statele Unite și aceea în care totul se bizuie în întregime pe puterea discreționară a judecătorilor – cum este cazul în Scoția.[14] În întreaga lume occidentală, regulile privind pronunțarea sentințelor tind să stabilească o sentință maximă (ca în Irlanda), o sentință minimă (ca în Canada) sau ambele (ca în Anglia și Țara Galilor)[15] și lasă la latitudinea judecătorilor ajustarea sentinței în plus sau în minus între aceste limite.

Niciun sistem nu este perfect. Există întotdeauna un amestec de nedreptăți care se iau la întrecere, un haos al injustițiilor care se opun una alteia. Dar în mijlocul acestui

conflict și cu toată această complexitate, un algoritm are prilejul să fie în elementul său. Pentru că – în mod remarcabil – cu un algoritm care ar face parte din proces se pot garanta atât consecvența, cât și justiția în fiecare caz în parte. Nimeni nu mai este nevoit să aleagă între cele două.

ECUAȚIA JUSTIȚIEI

Algoritmii nu pot lua decizii în ceea ce privește vinovăția cuiva. Ei nu pot să cântărească argumentele aduse de apărare și de acuzare, nici să studieze dovezile și nici să stabilească dacă un inculpat are într-adevăr remușcări. Prin urmare, nu trebuie să ne așteptăm ca aceștia să înlocuiască judecătorii în viitorul apropiat. Ceea ce poate să facă un algoritm, oricât de incredibil ar părea, este să proceseze informațiile despre o anumită persoană și să calculeze care este riscul ca aceasta să recidiveze. Iar, din moment ce deciziile judecătorilor se bazează pe probabilitatea ca un inculpat să comită din nou o infracțiune, aceasta se dovedește o aptitudine destul de utilă.

Informațiile și algoritmii se folosesc în sistemul judiciar de aproape un secol, primele exemple datând din America anilor 1920. Pe atunci, în sistemul american sentințele infractorilor condamnați aveau o durată maximă, iar, după o anumită perioadă, aceștia deveneau eligibili pentru eliberarea condiționată*. Zeci de mii de prizonieri au beneficiat astfel de eliberarea înainte de termen. Unii au fost reabilitați cu succes, alții nu. Însă, împreună, au constituit „baza de date" perfectă pentru organizarea unui experiment menit să descopere: este posibil să prezici dacă un deținut își va respecta sau nu condițiile eliberării condiționate?

* *Parole* (engl.). Un fapt amuzant: „parole" vine din francezul *parole*, care înseamnă „cuvinte vorbite". Se regăsește în forma originală prin anii 1700, când prizonierii erau eliberați dacă își dădeau cuvântul de onoare că nu vor recidiva. Sursa: https://www.etymonline.com/word/parole (n. a.).

Şi acum îşi face intrarea în scenă Ernest W. Burgess, un sociolog canadian de la Universitatea din Chicago pasionat de predicţii. Burgess a fost un mare susţinător al cuantificării fenomenelor sociale. În cursul carierei sale, a încercat să prognozeze totul – de la efectele pensionării şi până la reuşita în căsnicie –, iar în 1928 a devenit prima persoană care a conceput cu succes un instrument pentru calcularea riscului de apariţie a comportamentului infracţional, unul care se baza mai mult pe măsurători decât pe intuiţie.

Folosind diverse informaţii despre 3 000 de deţinuţi din trei închisori din statul Illinois, Burgess a identificat 21 de factori pe care i-a considerat „posibil semnificativi" pentru a determina şansele ca un deţinut să îşi încalce termenii eliberării condiţionate. Printre aceştia se numărau: genul infracţiunii, numărul de luni petrecute în închisoare şi tipul social al deţinutului, pe care – cu toată „delicateţea" la care ne putem aştepta din partea unui sociolog de la începutul secolului al XX-lea – i-a împărţit în categorii precum „vagabond", „beţiv", „om de nimic", „băiat de la fermă" şi „imigrant".[16]

Burgess i-a alocat fiecărui deţinut un punctaj între zero şi unu la fiecare dintre aceşti douăzeci şi unu de factori. Pe bărbaţii care obţineau punctaje mari (între 16 şi 21) îi considera a fi mai puţin predispuşi să recidiveze, pe aceia care aveau punctaje mici (patru sau mai puţin) îi considera susceptibili să îşi încalce condiţiile de eliberare condiţionată.

Când toţi condamnaţii au fost eliberaţi, fiind, prin urmare, liberi să încalce condiţiile de eliberare condiţionată dacă ar fi ales să facă acest lucru, Burgess a avut ocazia să verifice cât de precise fuseseră predicţiile sale. Cu ajutorul unei analize atât de simpliste, el a reuşit să fie remarcabil de exact. 98% dintre aceia aflaţi în grupul său cu risc scăzut au reuşit să depăşească cu bine perioada de eliberare condiţionată, iar două treimi dintre cei din grupul considerat ca având un risc ridicat au eşuat.[17] După cum s-a

dovedit, până și modelele statistice primitive au putut face predicții mai precise decât specialiștii.

Însă activitatea lui Burgess și-a avut criticii săi. Observatori sceptici s-au întrebat în ce măsură factorii care preziceau în mod fiabil succesul respectării condițiilor de eliberare condiționată într-un anumit loc și într-un anumit moment puteau fi aplicați în altă parte. (Aveau dreptate: nu sunt foarte convinsă că, de exemplu, categoria „băiat de la fermă" ar fi de mare ajutor în estimarea pericolului de recidivare printre infractorii urbani din zilele noastre.) Alți oameni de știință l-au criticat pe Burgess pentru că nu făcea decât să utilizeze toate informațiile de care dispunea, fără să cerceteze dacă acestea erau sau nu relevante.[18] S-au mai pus și întrebări referitoare la modul în care acorda el punctaje deținuților: în definitiv, metoda sa era ceva mai mult decât o opinie tradusă în ecuații. Nu este mai puțin adevărat că puterea lui de previziune a fost atât de impresionantă, încât, în 1935, metoda Burgess a pătruns în închisorile din statul Illinois pentru a ajuta comisiile de eliberare condiționată să ia decizii.[19] Iar pe la începutul secolului al XXI-lea, peste tot în lume se utilizau modele matematice care erau „urmașele" metodei lui Burgess.[20]

Revenind la epoca modernă, algoritmii de ultimă generație folosiți pentru evaluarea riscurilor în curțile de judecată sunt mult mai sofisticați decât instrumentele rudimentare inventate de Burgess. Aceștia nu vin numai în sprijinul deciziilor de eliberare condiționată, ci sunt întrebuințați și ca să ajute la repartizarea corectă a deținuților în diverse programe de intervenție, pentru a decide cui să i se ceară cauțiune și, de puțin timp, pentru a sprijini judecătorii în luarea deciziilor de condamnare. Principiul lor fundamental de funcționare este cel dintotdeauna: în sistem intră informațiile despre acuzat – vârstă, antecedente penale, gravitatea infracțiunii comise și așa mai departe – și iese o predicție referitoare la cât de mare ar fi riscul de a-l lăsa în libertate.

Aşadar, cum funcţionează totul? Ei bine, în mod general, algoritmii cei mai performanţi din prezent folosesc o tehnică numită *random forest* („pădure aleatorie"), care, în esenţă, are la bază o idee extrem de simplă: umilul arbore de decizie.*

ÎNTREABĂ PUBLICUL

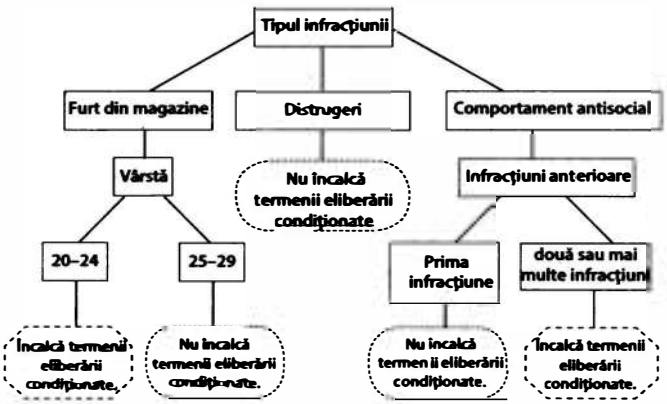
S-ar putea să fiţi familiarizaţi cu arborii de decizie din ceea ce aţi învăţat la şcoală. Ca metodă de structurare a observaţiilor, precum darea cu banul sau aruncarea zarurilor, aceştia sunt populari în rândul profesorilor de matematică. Odată construit, un arbore de decizie poate fi folosit ca o diagramă: se ia un set de împrejurări şi se analizează pas cu pas ce trebuie făcut sau, în cazul de faţă, ce se va întâmpla.

Imaginaţi-vă că încercaţi să hotărâţi dacă să cereţi cauţiune pentru o anumită persoană. Ca şi în cazul eliberării condiţionate, această decizie se bazează pe un calcul clar. Vinovăţia este irelevantă. Nu trebuie decât să faceţi o predicţie: va încălca inculpatul condiţiile eliberării pe cauţiune dacă i se va permite să iasă din închisoare?

Ca să vă ajute să faceţi o predicţie, dispuneţi de informaţiile culese de la un număr redus de recidivişti, dintre care unii au fugit sau au recidivat în timpul perioadei de cauţiune, iar alţii nu. Pe baza acestor informaţii, vă puteţi imagina că desenaţi un arbore de decizie simplu, cum este cel prezentat în continuare, folosind caracteristicile fiecărui inculpat pentru a construi o diagramă. Odată terminat, arborele de decizie este capabil să prevadă cum s-ar putea comporta un alt inculpat.

* *Decision tree* (engl.), una dintre metodele de *data mining* constând în reprezentarea unei serii de reguli care conduce la o valoare sau la o clasă de valori. Poate fi descris şi ca o diagramă care reprezintă variantele de acţiune şi posibilele rezultate pentru fiecare variantă alternativă (n. tr.).

Nu trebuie decât să urmăriți ramurile relevante în funcție de caracteristicile inculpatului, până când ajungeți la o predicție. Atâta vreme cât aceste caracteristici se încadrează în modelul oricui a fugit până în acel moment în timp ce se afla sub cauțiune, predicția va fi exactă.



Din acest moment însă arborii de decizie precum aceia pe care i-am construit în școală încep să dea greș. Pentru că, evident, nu toată lumea *urmează* tiparul celor care au fugit mai înainte. De sine stătător, arborele acesta va greși o mulțime de predicții. Și asta nu numai din cauză că începem cu un exemplu simplu. Chiar și cu un set de date uriaș despre cazurile anterioare și o diagramă îngrozitor de complicată care să îi corespundă, folosirea unui *singur* arbore de decizie ar putea fi doar cu puțin mai bună decât ghicitul la întâmplare.

În cazul în care construiești mai mult de un arbore, însă, totul se poate schimba. În loc să folosești toate informațiile dintr-odată, există o cale de a dezbina și de a cuceri. În ceea ce este cunoscut drept un *ensemble**, mai întâi, construiești mii

* „Ansamblu" (fr. în original) (n. tr.).

de arbori mai mici din subramificații aleatorii de informații. Apoi, când ți se prezintă un nou inculpat, îi ceri pur și simplu fiecărui arbore să decidă dacă el consideră că solicitarea cauțiunii este sau nu o idee bună. S-ar putea ca nu toți arborii să fie de acord în unanimitate, iar aceia de sine stătători încă ar putea să emită predicții nu tocmai valide, dar, dacă iei în considerare media tuturor răspunsurilor, vei putea să-ți îmbunătățești în mod spectaculos precizia predicțiilor.

Seamănă puțin cu momentul în care adresezi o întrebare publicului, în cadrul emisiunii *Vrei să fii miliardar?* O sală plină de necunoscuți va avea dreptate mai des decât cea mai inteligentă persoană pe care o cunoști. (Varianta ajutătoare „Întreabă publicul" are o rată de succes de 91%, prin comparație cu doar 65% a aceleia denumite „Sună un prieten".)[21] Greșelile făcute de un grup mai numeros se pot anula reciproc, ceea ce are drept rezultat că o mulțime devine mai înțeleaptă decât o singură persoană.

Același lucru este valabil și în cazul grupului mare de arbori de decizie, care, luați împreună, alcătuiesc o „pădure aleatorie", *random forest*, (jocul de cuvinte este intenționat). Pentru că predicțiile algoritmilor se bazează pe tiparele pe care aceștia le învață din informații, *random forest* este definit ca un algoritm de învățare automată care apare sub umbrela mai largă a inteligenței artificiale. (Sintagma „învățare automată" a apărut pentru prima oară în capitolul „Puterea", iar, mai departe, vom întâlni mai mulți algoritmi grupați sub această denumire generică; pentru moment, însă, merită menționat cât de spectaculos sună această descriere atunci când algoritmul este, în esență, diagrama pe care obișnuiam s-o desenăm la școală la care se adaugă un pic de prelucrare matematică.) *Random forests* s-au dovedit nemaipomenit de utili într-o multitudine de aplicații din lumea reală. Sunt utilizați de Netflix pentru a prezice ce ți-ar plăcea să urmărești în continuare pe baza preferințelor tale din trecut[22], de Airbnb ca să detecteze conturile

frauduloase[23] şi în sistemul de sănătate pentru diagnosticarea afecţiunilor (mai multe pe această temă, în capitolul următor).

Atunci când sunt folosiţi pentru evaluarea infractorilor, au două imense avantaje faţă de omologii lor umani. În primul rând, orice algoritm va da întotdeauna exact acelaşi răspuns atunci când i se va prezenta acelaşi set de circumstanţe. Consecvenţa este garantată, dar nu cu preţul unei decizii individualizate. Şi mai există un avantaj important: algoritmul va face predicţii mult mai bune.

LUPTA DINTRE OM ŞI MAŞINĂ

În anul 2017, mai mulţi cercetători şi-au propus să descopere cât de adecvate erau predicţiile unui calculator prin comparaţie cu deciziile luate de un anumit număr de judecători (umani).[24]

Pentru a-i ajuta în această misiune, cercetătorii aveau acces la dosarele fiecărei persoane arestate în oraşul New York pe o perioadă de cinci ani, între 2008 şi 2013. În acest interval de timp, trei sferturi dintr-un milion de persoane fuseseră audiate pentru a li se stabili cauţiunea, ceea ce a generat cu prisosinţă informaţii pentru a testa performanţele unui algoritm faţă de acelea ale unui judecător uman.

Sistemul judiciar din New York nu utilizase niciun algoritm în aceste cazuri, însă, după ce le-au analizat pe toate, retrospectiv, cercetătorii s-au apucat de treabă şi au construit o mulţime de arbori de decizie pentru a vedea cât de bine ar fi putut cineva să prevadă, când se derulaseră procesele, riscul ca inculpaţii să încalce prevederile eliberării condiţionate. Printre informaţiile despre infractori care au fost introduse se numărau: cazierul, delictul pe care tocmai îl comiseseră şi aşa mai departe. De aici rezulta probabilitatea ca inculpatul să încalce sau nu prevederile eliberării condiţionate.

În evidenţele reale, un număr de 408 283 de inculpaţi fuseseră eliberaţi pe cauţiune chiar înainte de proces. Oricare

dintre aceştia era liber să fugă sau să comită alte infracţiuni, ceea ce însemna că puteau fi folosite avantajele unei analize retrospective pentru a se verifica acurateţea predicţiilor algoritmilor şi pe aceea a deciziilor umane, deoarece se ştia cu exactitate cine nu se prezentase ulterior în faţa instanţei (15,2%) şi cine fusese arestat din nou pentru comiterea altei infracţiuni în perioada eliberării pe cauţiune (25,8%).

Din păcate pentru ştiinţă, oricărui inculpat considerat de către judecător ca având un risc ridicat i s-ar fi refuzat eliberarea pe cauţiune în acel moment – prin urmare, în acele situaţii nu se putea dovedi în niciun fel dacă evaluările judecătorilor fuseseră corecte sau greşite. Acest lucru complică situaţia. Înseamnă că este imposibil să obţinem o cifră concretă şi exactă care să surprindă cât de precişi au fost judecătorii per ansamblu. Iar în lipsa unui adevăr întemeiat pe fapte*, pe observaţii directe ale modului în care s-ar fi comportat inculpaţii care nu fuseseră eliberaţi, nu se poate stabili o acurateţe de ansamblu nici pentru algoritmi. În schimb, trebuie făcute deducţii avizate despre modul în care s-ar fi comportat infractorii rămaşi în închisoare dacă ar fi fost eliberaţi[25], iar comparaţiile dintre predicţiile oamenilor şi ale maşinilor trebuie realizate oarecum indirect.

Un lucru este, însă, sigur: judecătorii şi calculatorul nu au emis aceleaşi predicţii. Cercetătorii au arătat că mulţi dintre inculpaţii pe care algoritmul îi semnalase ca fiind băieţi foarte răi fuseseră trataţi de către judecători ca şi când ar fi prezentat un risc scăzut. De fapt, aproape jumătate dintre inculpaţii pe care algoritmul îi semnalase ca reprezentând grupul cu riscul cel mai mare au fost eliberaţi pe cauţiune de judecători.

Dar cine a avut dreptate? Informaţiile au arătat că grupul de care era preocupat algoritmul prezenta, într-adevăr, un risc. Puţin peste 56% din acest grup nu s-au mai prezentat în instanţă, iar 62,7% au comis noi infracţiuni în perioada

* *Ground truth* (engl.) (vezi glosar, p. 318) (n. red.).

eliberării condiționate – inclusiv pe cele mai abominabile: violul și crima. Algoritmul prevăzuse toate acestea.

Cercetătorii au susținut că, oricum ar fi fost folosit, algoritmul lor depășea cu mult un judecător uman în ceea ce privea performanțele. Iar cifrele le dădeau dreptate. Dacă dorești să încarcerezi mai puțini oameni în așteptarea procesului, algoritmul ar putea veni în sprijinul acestui fapt prin condamnarea la închisoare a 41,8% mai puțini inculpați, menținând totodată aceeași rată a criminalității. Sau, dacă ai fi mulțumit de procentul actual de inculpați eliberați pe cauțiune, și asta ar fi bine: doar fiind mai precis în selectarea inculpaților care urmează a fi eliberați condiționat, algoritmul ar putea să reducă procentul celor care fug și dispar în perioada cauțiunii cu 24,7%.

Aceste beneficii nu sunt doar teoretice. Statul american Rhode Island, unde instanțele au utilizat acest tip de algoritmi în ultimii opt ani, a reușit să obțină o reducere cu 17% a persoanelor încarcerate și o scădere de 6% a ratei de recidivă. Aceasta înseamnă că sute de infractori care prezintă un risc scăzut nu mai sunt ținuți în închisoare inutil și că sute de infracțiuni nu au mai fost comise. În plus, dat fiind că detenția unui prizonier costă peste 30 000 de lire sterline pe an în Marea Britanie[26] – pe când în Statele Unite petrecerea unui an într-o închisoare de maximă securitate poate să coste cam cât un an de studiu la Universitatea Harvard[27] –, aceasta înseamnă economisirea banilor a sute de mii de contribuabili. Este o situație din care toată lumea are de câștigat.

Sau poate că nu?

GĂSIREA LUI DARTH VADER

Firește că niciun algoritm nu poate să prevadă în mod *perfect* ce urmează să facă o persoană în viitor. Ființele umane sunt prea dezordonate, iraționale și impulsive pentru ca vreo

prognoză să fie vreodată cu totul *sigură* de ceea ce urmează să se întâmple. Algoritmii ar putea să ofere predicții mai bune, însă vor continua să facă greșeli. Întrebarea este: ce se întâmplă cu toți acei oameni al căror punctaj de risc a fost calculat greșit?

Algoritmul poate face două tipuri de greșeli. Richard Berk, profesor de criminologie și statistică la Universitatea din Pennsylvannia, un pionier în domeniul predicției recidivismului, are un mod remarcabil de a le descrie:

> Există băieți buni și băieți răi, mi-a spus el. Algoritmul tău întreabă, efectiv: „Care dintre ei sunt Darth Vader-i? Și care din ei sunt Luke Skywalker-i?"

Să-l lași liber pe Darth Vader este un tip de eroare cunoscută drept „fals negativ". Apare ori de câte ori nu reușești să identifici riscul pe care-l prezintă un anumit individ.

Pe de altă parte, să-l închizi pe Luke Skywalker ar fi un „fals pozitiv". Apare atunci când algoritmul identifică, în mod incorect, un individ ca fiind o persoană cu grad mare de risc.

Aceste două tipuri de erori, fals pozitiv și fals negativ, nu sunt specifice doar fenomenului de recidivă. Le vom întâlni, în mod repetat, pe parcursul acestei cărți. Orice algoritm care urmărește să clasifice ceva se poate face vinovat de comiterea acestor erori.

Algoritmii lui Berk susțin că pot să prezică, și cu o precizie de 75% încă, dacă cineva urmează să comită o omucidere, ceea ce îi face să fie algoritmii cei mai precișI din câți există.[28] Și, dacă ne gândim cât de liberă credem noi că este voința noastră, acest nivel de precizie este teribil de impresionant. Dar chiar și o precizie de 75% tot înseamnă că vor exista o mulțime de Luke Skywalker-i cărora le va fi refuzată eliberarea pe cauțiune din închisoare pentru că seamănă cu Darth Vader-ii de afară.

Consecințele etichetării greșite a unui inculpat devin încă și mai grave atunci când algoritmii sunt utilizați pentru a determina sentințele, în loc să se limiteze la stabilirea cauțiunii

sau a eliberării condiționate. Aceasta este o realitate modernă: de curând, unele state din SUA au început să le permită judecătorilor să vadă scorul de risc al unui infractor condamnat în timp ce deliberează asupra duratei întemnițării sale. Este o evoluție care a stârnit o dezbatere aprinsă, și nu fără motiv: una este să calculezi dacă să eliberezi sau nu o persoană mai devreme din închisoare și este cu totul altceva să calculezi de la bun început perioada în care aceasta să fie încarcerată.

O parte a problemei este că a decide durata unei sentințe înseamnă a lua în considerare mult mai mulți factori decât riscul ca un infractor să recidiveze – ceea ce reprezintă cam tot sprijinul pe care-l poate oferi un algoritm. Un judecător mai trebuie să aprecieze și riscul pe care infractorul îl reprezintă pentru alții, efectul de descurajare pe care hotărârea de condamnare îl va avea asupra altor infractori, problema despăgubirilor acordate victimei și șansele de reabilitare ale inculpatului. Sunt o mulțime de lucruri de pus în balanță, așa încât nu este de mirare că oamenii obiectează împotriva acordării unei importanțe atât de mari unui algoritm în luarea acestui tip de decizie. Și nici nu este de mirare că oamenilor li se par atât de profund tulburătoare întâmplări precum aceea prin care a trecut Paul Zilly.[29]

Acesta a fost condamnat pentru furtul unei mașini de tuns iarba. În februarie 2013, s-a prezentat în fața judecătorului Babler din Baron County, statul Wisconsin, știind că avocatul apărării ajunsese deja la o înțelegere cu acuzarea. Ambele părți căzuseră de acord că, în cazul său, o condamnare pe termen lung nu era cea mai bună soluție. Zilly se aștepta ca judecătorul să confirme, pur și simplu, înțelegerea aceasta.

Din păcate pentru el, judecătorii din statul Wisconsin utilizau un algoritm de evaluare a riscului brevetat, denumit COMPAS. Ca și în cazul instrumentului bugetar din Idaho menționat la capitolul „Puterea", funcționarea internă a COMPAS-ului era considerată un secret comercial. Spre deosebire de instrumentul bugetar, însă, codul algoritmului

COMPAS nu este accesibil publicului nici în prezent. Totuşi, din câte ştim, calculele se bazează pe răspunsurile pe care le dă un inculpat atunci când completează un chestionar. Acesta conţine întrebări precum: „Un om flămând are dreptul să fure; eşti de acord sau nu eşti de acord?" şi „Dacă ai locuit împreună cu ambii părinţi, iar aceştia s-au despărţit, ce vârstă aveai în acel moment?"[30] Algoritmul fusese conceput cu unicul scop de a prezice cât de probabil era ca un inculpat să recidiveze în următorii doi ani, iar în această activitate obţinuse un grad de precizie de circa 70%.[31] Altfel spus, s-ar fi înşelat, aproximativ, în legătură cu unul din trei inculpaţi. Cu toate acestea, el era folosit de judecători pentru stabilirea sentinţelor.

Punctajul lui Zilly nu era prea bun. Algoritmul îl evaluase ca prezentând risc ridicat de a comite infracţiuni violente în viitor şi risc mediu de recidivă per ansamblu: „Când verific evaluarea riscurilor, spusese judecătorul Babler în instanţă, aceasta vă este cât se poate de nefavorabilă."

După ce văzuse punctajul lui Zilly, judecătorul a dat dovadă de mai multă încredere în algoritm decât în înţelegerea la care ajunseseră apărarea şi acuzarea, aşa că a respins-o şi i-a dublat sentinţa lui Zilly de la un an petrecut în închisoarea districtuală, la doi într-o închisoare de stat, mai restrictivă.

Este imposibil să ştim cu certitudine dacă Zilly merita punctajul obţinut pentru risc ridicat, cu toate că un procent de acurateţe de 70% pare un prag destul de scăzut pentru a justifica utilizarea algoritmului în vederea respingerii altor factori.

Cazul lui Zilly a fost larg mediatizat, însă nu este singurul exemplu de acest gen. În 2003, Christopher Drew Brooks, un tânăr de 19 ani, a întreţinut relaţii sexuale consensuale cu o fată de 14 ani şi a fost acuzat de întreţinerea unor raporturi sexuale cu un minor de o instanţă de judecată din statul Virginia. Iniţial, instrucţiunile cu privire la sentinţe sugerau o perioadă de încarcerare între şapte şi 16 luni pentru această

faptă, însă după ce recomandările au fost modificate pentru a include și punctajul de risc (care nu fusese calculat de COMPAS, în cazul lui), limita superioară a fost majorată la 24 de luni. Luând în calcul această informație, judecătorul l-a condamnat pe Christopher la 18 luni de închisoare.[32]

Iată unde este problema. Algoritmul folosea vârsta ca factor în calcularea punctajului de recidivă. Faptul că fusese condamnat pentru un delict sexual la o vârstă atât de fragedă era în defavoarea lui Brooks, chiar dacă asta însemna totuși că vârsta îi era mai apropiată de cea a victimei. De fapt, dacă Brooks ar fi avut 36 de ani (prin urmare, cu 22 de ani mai mult decât fata), algoritmul ar fi recomandat să nu fie trimis deloc la închisoare![33]

Acestea nu sunt singurele exemple în care oamenii au avut mai multă încredere în rezultatele obținute de un calculator decât în propria putere de judecată și nu vor fi nici ultimele. Se pune întrebarea: ce este de făcut în această privință? Curtea Supremă din Wisconsin vine cu o sugestie. Referindu-se, în mod special, la pericolul ca judecătorii să se bazeze prea mult pe algoritmul COMPAS, aceasta a decretat:

> Ne așteptăm ca tribunalele să manifeste discernământ, în privința fiecărui inculpat în parte, atunci când vor evalua vreunul dintre punctajele de risc obținute cu ajutorul lui COMPAS.[34]

Richard Berk sugerează însă că aceasta ar putea fi o abordare prea optimistă:

> Judecătoriile sunt preocupate să nu facă greșeli – în special, judecătorii numiți din oficiu. Algoritmul le oferă o modalitate de a-și micșora volumul de muncă și, totodată, de a nu fi considerați răspunzători pentru o anumită sentință.[35]

Dar mai este o problemă. Dacă un algoritm clasifică o persoană ca prezentând risc ridicat, iar judecătorul refuză, în consecință, punerea ei în libertate, nu există nicio cale de a ști cu certitudine dacă algoritmul i-a analizat perspectivele cu acuratețe. Să-l luăm drept exemplu pe Zilly. Poate că ar fi

continuat să aibă un comportament violent. Poate că nu. Poate că a fi considerat un condamnat cu nivel de risc ridicat și a fi trimis într-o închisoare de stat nu au făcut decât să-i impună o cu totul altă traiectorie în viață decât dacă s-ar fi ținut cont de înțelegerea la care se ajunsese la tribunal. Neavând nicio modalitate de a verifica predicțiile algoritmilor, nu putem ști dacă judecătorul a avut dreptate să dea crezare punctajului pentru risc și nu avem nicio cale de a verifica dacă Zilly era, de fapt, un Darth Vader sau un Luke Skywalker.

Această problemă nu are o soluție facilă. Cum îi convingi pe oameni să apeleze la o doză sănătoasă de bun-simț când se pune problema utilizării acestor algoritmi? Și, chiar dacă acest lucru ar fi posibil, mai apare o problemă legată de predicția recidivei. Fără îndoială, cea mai controversată dintre toate.

PREJUDECĂȚILE MAȘINILOR

În anul 2016, redacția publicației independente on-line ProPublica, prima care a relatat cazul lui Zilly, a studiat în detaliu algoritmul COMPAS și a reconstituit în sens invers predicțiile pe care le făcuse acesta despre perspectivele de recidivare a peste 7 000 de infractori reali din Florida[36], în cazuri datând din 2013 până în 2014. Cercetătorii doreau să verifice acuratețea punctajelor obținute de COMPAS, urmărind care dintre infractori recidivaseră efectiv. Mai doreau, de asemenea, să vadă dacă exista vreo diferență între riscurile prezise pentru infractorii de culoare și cei albi.

Cu toate că algoritmul nu include, în mod explicit, rasa ca factor, jurnaliștii au descoperit că nu toți inculpații erau tratați în mod egal la efectuarea calculelor. Deși șansele de eroare ale algoritmului erau, în linii mari, aceleași pentru infractorii de culoare și pentru cei albi, el făcea diverse tipuri de erori în cazul fiecărui grup rasial în parte.

Dacă erai unul dintre infractorii care nu mai comiseseră nicio faptă reprobabilă după arestarea inițială, prin urmare,

un adevărat Luke Skywalker, algoritmul era de două ori mai predispus să te eticheteze în mod greşit ca având un risc ridicat dacă erai o persoană de culoare, decât dacă erai alb. Rezultatele fals pozitive ale algoritmului erau în mod disproporţionat atribuite persoanelor de culoare. În schimb, printre infractorii care au comis o nouă infracţiune în următorii doi ani, printre Darth Vadder-i, adică, era de două ori mai probabil ca deţinuţii albi să fie evaluaţi de algoritm ca prezentând un risc scăzut pentru societate, faţă de deţinuţii de culoare. Rezultatele fals negative ale algoritmului erau, în mod disproporţionat, atribuite albilor.

Deloc surprinzător, analiza făcută de ProPublica a stârnit indignare pe tot cuprinsul Statelor Unite şi chiar mai departe. S-au scris sute de articole care exprimau dezaprobarea fermă a autorilor faţă de utilizarea unor calcule impersonale în procesul uman de a face dreptate şi se condamna folosirea unor algoritmi imperfecţi, părtinitori, în decizii care aveau un impact atât de mare asupra viitorului unei persoane. Multe dintre aceste critici sunt greu de contrazis – toată lumea merită să aibă parte de un tratament corect şi egal indiferent cine evaluează cazul respectiv –, iar studiul realizat de ProPublica nu pune algoritmul într-o lumină prea bună.

Să fim, însă, atenţi, pentru o clipă, la tendinţa noastră de a înlătura „algoritmii imperfecţi". Înainte de a respinge cu totul utilizarea acestora în sistemul judiciar, merită să ne punem întrebarea: cum v-aţi aştepta să fie un algoritm nepărtinitor?

Aţi dori, desigur, ca acesta să facă predicţii la fel de corecte pentru condamnaţii de culoare şi pentru cei albi. Pare, de asemenea, rezonabil să ceri ca toţi factorii ce se consideră a reprezenta un „risc ridicat" să fie la fel pentru toată lumea. Algoritmul ar trebui să fie la fel de performant în selectarea infractorilor predispuşi să recidiveze indiferent de grupul rasial (sau oricare altul) căruia îi aparţin aceştia. La care, aşa cum a subliniat ProPublica, se adaugă faptul că algoritmul ar

trebui să facă același tip de greșeli cu aceeași frecvență pentru toată lumea – indiferent de rasă.

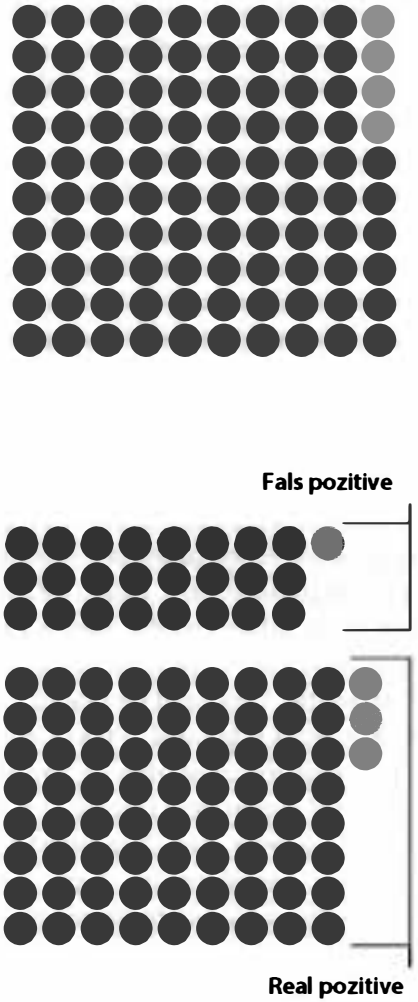
Niciuna dintre aceste patru afirmații nu pare să fie din calea-afară de ambițioasă. Numai că apare, totuși, o problemă. Din nefericire, unele tipuri de corectitudine sunt, din punct de vedere matematic, incompatibile cu altele.

Dați-mi voie să vă lămuresc. Imaginați-vă că opriți oamenii pe stradă și utilizați un algoritm care prezice dacă o anumită persoană urmează să comită un omor. Însă, din moment ce marea majoritate a omorurilor sunt comise de bărbați (în întreaga lume, 96% dintre ucigași sunt bărbați)[37], dacă algoritmul care depistează ucigași trebuie să facă predicții exacte, el va identifica, în mod necesar, mai mulți bărbați decât femei ca prezentând un risc mai mare de a comite acest gen de faptă.

Să presupunem că algoritmul nostru de detectare a ucigașilor are o rată de predicție de 75%. Asta înseamnă că trei sferturi din persoanele pe care algoritmul le consideră ca având un risc crescut sunt, într-adevăr, niște Darth Vader-i.

În cele din urmă, după ce ați oprit destui necunoscuți, veți avea 100 de persoane semnalate de algoritm ca fiind potențiali ucigași. Pentru a face corelația cu statistica făptașilor, 96 dintre acești 100 vor fi, obligatoriu, bărbați. Patru vor fi femei. În continuare este o poză pentru ilustrare. Bărbații sunt reprezentați de cercuțele închise la culoare, femeile apar ca niște cercuțe de un gri-deschis.

Acum, din moment ce algoritmul face predicții corecte atât pentru bărbați, cât și pentru femei într-o proporție egală de 75%, un sfert dintre femei și un sfert din bărbați vor fi, de fapt, Luke Skywalker-i: persoane care au fost incorect identificate ca având un nivel de risc ridicat, atunci când, de fapt, ele nu prezintă niciunul.



Când faci calculele, aşa cum se poate vedea din a doua imagine de mai sus, vor fi acuzaţi pe nedrept mai mulţi bărbaţi nevinovaţi decât femei nevinovate doar în virtutea faptului că bărbaţii săvârşesc mai multe omoruri decât femeile.

Asta nu are nimic de-a face nici cu crima în sine şi nici cu algoritmul: este doar o certitudine matematică. Rezultatul este viciat de prejudecată pentru că realitatea însăşi este deformată într-un mod care întăreşte această prejudecată. Mai mulţi bărbaţi comit omucideri, aşadar, mai mulţi bărbaţi vor fi acuzaţi, în mod greşit, că prezintă riscul de a ucide.*

Dacă nu se întâmplă ca procentul persoanelor care comit crime să fie acelaşi în fiecare grup de infractori, este matematic imposibil să concepi un test care, în acelaşi timp, să facă predicţii tot atât de exacte pe toată linia şi să eticheteze persoanele din fiecare grup de infractori drept fals pozitive şi fals negative în aceeaşi proporţie.

Desigur, persoanele de culoare din SUA au făcut obiectul unor prejudecăţi şi nedreptăţi cumplite de-a lungul secolelor. Din această cauză, ele continuă să apară disproporţionat de des la baza piramidei socio-economice şi la nivelurile superioare ale statisticilor privind criminalitatea. Există, de asemenea, dovezi că – pentru anumite infracţiuni din Statele Unite cel puţin – populaţia de culoare este într-un număr disproporţionat de mare vizată de poliţie. Consumul de marijuana, de exemplu, se întâlneşte în aceeaşi măsură şi la persoanele de culoare şi la cele albe, însă numărul arestărilor de afro-americani poate fi chiar şi de opt ori mai mare.[38] Oricare ar fi motivul acestei discrepanţe, tristul adevăr este că în Statele Unite numărul arestărilor nu este acelaşi pentru toate grupurile rasiale. Persoanele de culoare sunt arestate din nou mai des decât albii. Algoritmul nu le discerne după culoarea pielii, ci după efectele mult prea previzibile ale vieţuirii în societatea americană, care, din motive istorice, este atât de dezechilibrată. Până când toate grupurile rasiale vor ajunge să fie arestate în proporţii egale,

* Un rezultat ca acesta poate surveni chiar dacă nu este folosit, în mod explicit, genul cuiva drept factor al algoritmului. Atât timp cât predicţia se bazează pe factori care se corelează mai mult cu un grup decât cu altul (cum ar fi antecedentele de manifestări violente ale acuzatului), acest tip de inechitate poate să apară (n. a.).

acest gen de inechitate va rămâne, în continuare, o certitudine matematică.

Asta nu înseamnă însă nici că ar trebui să respingem cu totul articolul din ProPublica. Analiza pe care o face acesta subliniază cât de ușor pot algoritmii să perpetueze inechitățile trecutului. Nu înseamnă însă nici că ar fi cazul să găsim scuze algoritmului COMPAS. Orice companie care obține un profit din analizarea informațiilor despre diverse persoane are responsabilitatea morală (dacă nu cumva și legală) să-și recunoască defectele și punctele vulnerabile. Equivant (fosta Northpointe), în schimb, firma care produce COMPAS, continuă să considere configurația algoritmului drept un secret care trebuie păzit îndeaproape pentru a-și proteja drepturile de proprietate intelectuală asupra acestuia.[39]

Avem totuși câteva opțiuni. Nu există nimic intrinsec în acești algoritmi care să îi determine să repete prejudecățile trecutului. Totul se rezumă la informațiile care le sunt furnizate. Putem alege să fim „adepții unui empirism cras" (cum spune Richard Berk) și să ne luăm, în continuare, după cifrele deja existente sau putem decide că statu-quoul este nedrept și, în consecință, să modificăm numerele statisticilor.

Ca să ofer o analogie, o să vă rog să încercați să dați o căutare pe Google Imagini pentru „profesor de matematică". Poate că nu veți fi surprinși să descoperiți că majoritatea imaginilor reprezintă bărbați albi, de vârstă mijlocie, stând în picioare în fața unor table negre, pline de însemne făcute cu creta. Căutarea mea a întors numai o singură imagine cu o femeie în primele 20 de imagini, ceea ce este o reflectare deprimant de corectă a realității: în jur de 94% dintre profesorii de matematică sunt bărbați.[40] Dar, oricât de precise ar fi rezultatele, se poate afirma că să utilizezi algoritmii doar ca pe o oglindă menită să reflecte lumea reală nu este întotdeauna util, mai ales atunci când oglinda reflectă o realitate prezentă care există numai din cauza unor îndelungate secole de nedreptăți. Dacă ar dori,

Google ar putea să-și ajusteze subtil algoritmul, astfel încât să redea, în mod prioritar, imagini cu femei sau cu profesori care să fie de altă rasă, nu doar din cea albă, pentru a echilibra astfel puțin balanța și a ajunge să reflecte societatea spre care tindem, mai degrabă decât pe aceea în care trăim.

Același lucru se întâmplă și în sistemul judiciar. Atunci când utilizăm un algoritm, s-ar cuveni să ne întrebăm, efectiv: ce procent dintr-un anumit grup ne-am aștepta să prezinte un risc ridicat într-o societate perfect echitabilă? Algoritmul ne oferă opțiunea să sărim direct la cifra aceea. Sau, dacă decidem că îndepărtarea dintr-odată a tuturor prejudecăților din sistemul judiciar nu este oportună, am putea cere, în schimb, algoritmului să acționeze progresiv, în timp, spre îndeplinirea acestui deziderat.

Dispunem de opțiuni și în ceea ce privește modul în care îi tratăm pe inculpații care prezintă un risc ridicat. La eliberarea pe cauțiune, unde riscul ca un inculpat să nu-și mai facă apariția la următorul termen al procesului este o componentă principală a predicției algoritmului, abordarea standard este aceea de a refuza acordarea cauțiunii oricărei persoane care prezintă un grad ridicat de risc. Dar algoritmul ar putea oferi și prilejul de a afla de ce o persoană ar putea să nu se prezinte la o înfățișare la tribunal. Are acces la un mijloc de transport adecvat ca să ajungă acolo? Are probleme cu îngrijirea copiilor pe care nu îi poate lăsa nimănui în grijă, ceea ce o s-o împiedice să se prezinte în instanță? Există nedreptăți sociale pe care algoritmul ar putea fi programat să le atenueze, în loc să le exacerbeze?

Răspunsurile la aceste întrebări ar trebui să vină din partea forumurilor deschise de dezbateri publice și din sferele guvernării, mai degrabă decât din sălile de ședință ale companiilor private. Din fericire, pentru a controla această industrie se solicită din ce în ce mai insistent apariția unui organism de reglementare a algoritmilor. Așa cum procedează US Food and

Drug Administration* în cazul produselor farmaceutice, acest organism ar putea testa, în spatele uşilor închise, acurateţea, consecvenţa şi distorsiunile algoritmilor şi ar avea autoritatea de a aproba sau de a respinge utilizarea unui produs care ar urma să afecteze fiinţele umane reale. Până atunci, însă, este extraordinar de important ca organizaţii cum este ProPublica să continue să tragă la răspundere companiile care concep algoritmi, atât timp cât acuzaţiile că aceştia au fost părtinitori nu se transformă în solicitări de a-i interzice complet, cel puţin, nu fără a cerceta cu atenţie ce ne-ar mai rămâne în acest caz.

DECIZII DIFICILE

Există un aspect foarte important la care trebuie să mă refer. Dacă respingem algoritmii, cu ce fel de sistem judiciar am mai rămâne? Pentru că inconsecvenţa nu este singurul defect de care s-a dovedit că se fac vinovaţi judecătorii.

Din punct de vedere legal, rasa, genul şi clasa socială a unui inculpat nu ar trebui să influenţeze niciodată nicio hotărâre judecătorească. (În definitiv, se presupune că justiţia este oarbă, nu?) Cu toate acestea, cei mai mulţi dintre judecători doresc să fie cât mai nepărtinitori cu putinţă, dovezile au arătat în mod repetat că judecătorii, efectiv, discriminează. Studii din Statele Unite au dovedit că acuzaţii de culoare vor fi trimişi la închisoare, în medie, un timp mai îndelungat[41], că au şanse mai mici să li se acorde eliberarea pe cauţiune[42], dar şanse mai mari să fie condamnaţi la moarte[43] şi, odată ce se află pe lista condamnaţilor la moarte, există o mai probabilitate să fie executaţi.[44] Alte cercetări au arătat că,

* Cunoscută şi cu acronimul FDA, este una dintre agenţiile federale din Statele Unite din domeniul medical. Este responsabilă cu reglementarea produselor alimentare (destinate oamenilor şi animalelor), a suplimentelor dietetice, medicamentelor, cosmeticelor, dispozitivelor medicale şi a acelora care emit radiaţii etc. (n. tr.).

pentru aceeași infracțiune, bărbații sunt tratați cu mai multă severitate decât femeile[45] și că inculpaților cu niveluri de venit și de educație scăzute li se dau condamnări mai lungi.[46]

Ca și în cazul algoritmilor, nu atât prejudecățile propriu-zise generează aceste rezultate părtinitoare, cât faptul că istoria se repetă. Prejudecățile sociale și culturale pot să apară, pur și simplu, ca o consecință automată a modului în care oamenii iau deciziile.

Pentru a explica de ce se întâmplă acest lucru, trebuie să înțelegem, mai întâi, câteva lucruri simple despre intuiția umană, așa că vom părăsi, pentru moment, sala de judecată, cât timp reflectați la următoarea problemă:

> O bâtă de baseball și o minge costă împreună 1,10 lire.
> Bâta costă cu o liră mai mult decât mingea.
> Cât costă mingea?

Această problemă, care a fost introdusă în best-sellerul economistului și psihologului Daniel Kahneman, laureat al Premiului Nobel pentru Economie, *Gândire rapidă, gândire lentă*[47], ilustrează o capcană importantă în care cădem cu toții atunci când gândim.

Întrebarea are un răspuns incorect, ușor de sesizat la o gândire mai atentă, dar și un răspuns incorect, care se ivește imediat în minte. Răspunsul la care v-ați gândit, mai întâi, a fost, bineînțeles, 10 pence.*

Să nu vă pară rău dacă nu ați găsit răspunsul corect (5 penny). Nu l-au găsit nici 71,8% din judecători atunci când au fost întrebați.[48] Chiar și cei care, până la urmă, reușesc să găsească răspunsul corect, trebuie să-și reprime, pe moment, impulsul de a-și urma intuiția.

Competiția dintre intuiție și gândirea chibzuită este esențială în cazul relatării noastre despre hotărârile judecătorești. Psihologii sunt de acord, în linii mari, că

* Dacă mingea ar costa 10 penny ar însemna că bâta a costat 1,10 lire, iar prețurile celor două produse, însumate, ar fi 1,20 lire (n. a.).

dispunem de două moduri de gândire. Primul sistem este automat, instinctiv, dar predispus la greşeli. (Este sistemul responsabil de răspunsul „10 penny" care ni s-a ivit în minte la întrebările de mai sus.) Al doilea sistem este încet, analitic, chibzuit, însă, adesea, destul de leneş.[49]

Dacă întrebi o persoană cum a ajuns să ia o decizie, al doilea sistem va formula un răspuns, dar, aşa cum spune Daniel Kahneman: „acesta vine adesea în sprijinul sau justifică ideile şi sentimentele generate de primul sistem".[50]

În această situaţie, judecătorii nu diferă de noi, ceilalţi. În definitiv, şi ei sunt oameni, cu aceleaşi toane şi slăbiciuni pe care le avem cu toţii. Adevărul este că minţile noastre pur şi simplu nu sunt echipate pentru evaluarea solidă şi raţională a problemelor mari şi complicate. Nu putem cântări cu uşurinţă diverşii factori ai unei situaţii şi să legăm totul de o manieră logică, în timp ce împiedicăm primul sistem intuitiv să intervină şi să o ia pe tot soiul de scurtături cognitive.

Când ne referim la justiţie, de exemplu, s-ar putea crede că judecătorii sunt capabili să evalueze un caz în totalitate, cântărind cu atenţie argumentele pro şi contra înainte de a lua o hotărâre. Din păcate, dovezile indică altceva. Psihologii au arătat, în schimb, că toată strategia folosită de judecători se rezumă la a trece în revistă o listă de semnale de avertizare pe care şi-au fixat-o în memorie. Dacă relatarea acuzatului îi evocă oricare dintre aceste semnale – condamnări din trecut, absenţa legăturilor cu comunitatea, solicitarea procurorului –, judecătorul se va opri din chibzuire şi îi va refuza eliberarea pe cauţiune.[51]

Problema este că foarte multe dintre aceste semnale de avertizare au legătură cu rasa, genul şi nivelul de educaţie al inculpatului. Judecătorii nu se pot abţine să nu se bazeze pe intuiţie mai mult decât ar trebui; iar procedând în felul acesta, ei perpetuează fără să vrea inechităţile sistemului.

Şi asta nu-i tot! Din păcate, abia dacă am atins o mică parte a problemei, abia dacă am deschis discuţia despre cât de jalnici, cât de incapabili sunt oamenii să se comporte ca nişte judecători corecţi şi nepărtinitori.

Dacă v-aţi convins vreodată singuri că un anumit obiect de îmbrăcăminte extrem de costisitor a devenit o afacere grozavă doar pentru că era vândut cu 50% mai ieftin (aşa cum eu, una, fac de obicei), atunci ştiţi tot ce e de ştiut despre aşa-numitul „efect de ancorare". Acesta denumeşte faptul că ne este greu să stabilim o valoare numerică pentru diverse lucruri şi găsim că este mult mai simplu să facem comparaţii între diverse valori, decât să găsim din senin o singură valoare. Comercianţii folosesc efectul de ancorare de ani de zile nu numai ca să influenţeze cât de mult apreciem anumite produse, dar şi ca să controleze cantitatea de produse pe care o cumpărăm. Gândiţi-vă, spre exemplu, la acele etichete indicatoare din supermarketuri pe care scrie: „Nu pot fi cumpărate mai mult de 12 cutii cu supă de către un client". Ele nu sunt menite să împiedice iubitorii de supă să epuizeze stocul de produse, aşa cum aţi putea crede. Dimpotrivă, se află acolo pentru a vă manipula, în mod subtil, percepţia asupra numărului de cutii cu supă de care aveţi nevoie. Creierul se ancorează de numărul 12 şi ajustează în jos. Iar un studiu din anii 1990 a arătat că un astfel de semnal ar putea să crească media vânzărilor per client de la 3,3 cutii cu supă, la 7.[52]

Acum nu veţi mai fi deloc surprinşi să aflaţi că şi judecătorii sunt sensibili la efectul de ancorare. Ei sunt mai predispuşi să ceară despăgubiri mai însemnate dacă procurorul solicită o sumă mai mare[53] şi să pronunţe o condamnare mai lungă dacă procurorul cere o pedeapsă mai severă.[54] Un studiu a arătat chiar că s-ar putea influenţa în mod considerabil durata unei condamnări, într-un caz ipotetic, dacă ar fi trimis un jurnalist la judecător în timpul unei pauze, iar acesta ar introduce, cu subtililitate, în conversaţie sugerarea unei anumite sentinţe.

(„Credeți că sentința, în cazul de față, ar trebui să fie condamnarea pe o durată mai mare sau mică de trei ani?")[55] Probabil că lucrul cel mai îngrozitor este că poți influența decizia unui judecător dacă îi ceri, pur și simplu, să dea cu banul înainte de a intra în sala de judecată.[56] Până și judecătorii cu cea mai mare experiență au fost susceptibili la acest gen de manipulare.[57]

Și mai există o fisură în modul în care ființele umane fac comparații între numere, între cantități, care are un anumit impact asupra caracterului echitabil al pronunțării sentinței. Poate ați remarcat deja această ciudățenie a minții omenești. Atunci când muzica e dată tare, creșterea volumului cu o unitate este mai puțin percepută decât aceeași creștere absolută la un volum mai mic; o creștere a prețului de la 1,20 lire sterline la 2,20 pare exagerat, însă o creștere de la 67 de lire sterline la 68 nu prea contează. Iar timpul pare să treacă mai repede pe măsură ce înaintăm în vârstă. Toate acestea se întâmplă pentru că la ființele umane simțurile funcționează mai mult în termeni relativi decât în valori absolute. Nu percepem fiecare an ca pe un interval de timp fix, ci îl sesizăm ca pe o fracțiune din ce în ce mai scurtă din viața pe care am trăit-o. Dimensiunea fragmentelor de timp, bani sau volum sonor pe care le percepem urmează o expresie matematică foarte simplă, cunoscută drept Legea lui Weber.

Pe scurt, Legea lui Weber afirmă că întotdeauna cea mai mică modificare a stimulului care poate fi percepută, așa-numita „cea mai mică diferență perceptibilă", este proporțională cu stimulul inițial. Deloc surprinzător, această descoperire a fost exploatată și de comercianți. Ei știu exact cât de mult pot să micșoreze o tabletă de ciocolată fără ca niciunul dintre clienți să bage de seamă sau știu precis cât de mult pot să majoreze prețul unui produs înainte ca vreunul dintre noi să considere că ar fi mai bine să-l caute și în alt magazin.

În contextul justiției, problema este că Legea lui Weber influențează durata condamnării alese de judecători. Diferențele dintre sentințe sunt tot mai mari, pe măsură ce

pedepsele devin mai severe. Dacă o infracțiune este cu puțin mai gravă decât una care merită o condamnare de 20 de ani, atunci trei luni în plus, să spunem, nu par suficiente: parcă nu ar exista o diferență suficient de mare între o durată a sentinței de 20 de ani și una de 20 de ani și trei luni. Dar bineînțeles că există: trei luni de închisoare sunt, totuși, trei luni de închisoare, indiferent de ce a fost înaintea lor. Cu toate acestea, în loc să mai adauge doar câteva luni, judecătorii vor trece la următoarea diferență perceptibilă a duratei sentinței, care, în cazul de față, este de 25 de ani.[58] Știm că acest lucru se întâmplă pentru că putem să comparăm duratele sentințelor pronunțate în realitate cu cele pe care le-ar prezice Legea lui Weber. Un studiu din 2017 a analizat peste 100 000 de sentințe, atât din Marea Britanie, cât și din Australia, și a găsit că până la 99% dintre inculpații considerați vinovați au primit o condamnare care corespundea acestui enunț.[59]

Mandeep Dhami, autorul principal al studiului, mi-a declarat:

> Nu are importanță ce tip de infracțiune ai comis sau ce fel de inculpat ești, nici în ce țară ai fost condamnat și nici chiar dacă ai fost condamnat la închisoare sau la prestarea unor ore de muncă în beneficiul societății.

Tot ce conta era numărul care se ivea în mintea judecătorului și care i se părea acestuia cât de cât corect.

Iar dacă ar fi să mă refer la cât de părtinitori pot fi judecătorii, din păcate, mi-ar fi posibil să continui la nesfârșit! Judecătorii care au fiice sunt mai predispuși să ia decizii favorabile femeilor.[60] Judecătorii sunt mai puțin dispuși să ofere eliberare pe cauțiune atunci când echipa sportivă locală a pierdut recent un meci. Iar un studiu celebru sugerează chiar că până și momentul din zi îți afectează șansele de a obține un rezultat favorabil în instanță.[61] Deși studiul acesta încă nu a fost reprodus[62] și sunt în curs câteva dezbateri în privința amplorii efectului, se pare că ar exista dovezi că susținerea unei depoziții chiar înainte de masa de prânz poate fi un dezavantaj: judecătorii din studiul inițial

erau mai dispuși să ofere eliberarea pe cauțiune dacă tocmai se întorseseră după o pauză și cel mai puțin dispuși atunci când se apropia pauza de masă.

Un alt studiu a arătat că un singur judecător va evita să pronunțe prea multe sentințe asemănătoare la rând. Prin urmare, șansele de a ți se oferi eliberarea pe cauțiune scad amețitor dacă patru cazuri care au fost judecate chiar înaintea ta au avut succes în această privință.[63]

Unii cercetători mai susțin că până și modul în care percepem persoanele necunoscute se schimbă în funcție de temperatura băuturii pe care o ținem în mână. Astfel, dacă ne-a fost oferită o băutură caldă chiar înainte de a întâlni o persoană nouă, sugerează ei, suntem mai dispuși să considerăm că aceasta are o personalitate mai caldă, mai generoasă, mai plină de afecțiune.[64]

Lista aceasta lungă conține numai chestiunile care pot fi măsurate. Fără îndoială că există nenumărați alți factori care ne influențează subtil comportamentul și care nu pot fi testați într-o sală de judecată.

ÎN CONCLUZIE

Voi fi sinceră cu dumneavoastră. Când am aflat pentru prima oară că algoritmii erau utilizați în instanțele de judecată, m-am opus acestei idei. Un algoritm va face greșeli, iar când o eroare poate însemna că un om își va pierde dreptul la libertate, nu mi se părea un gest responsabil să lași această putere la latitudinea unei mașini.

Nu sunt singura care gândește astfel. Multe persoane (poate chiar majoritatea) care ajung să fie la cheremul sistemului de justiție penală sunt de aceeași părere. Mandeep Dhami mi-a spus ce simțeau inculpații cu care lucrase față de modul în care erau luate deciziile privitoare la viitorul lor:

> Chiar dacă știau că judecătorul ar fi putut face mai multe greșeli, inculpații continuau să prefere să fie judecați de o ființă umană, nu de un algoritm. Își doreau să existe o relație umană.

La fel și avocații, dacă tot a venit vorba. Un avocat al apărării care își practică meseria la Londra, cu care am vorbit, mi-a spus că rolul său în sala de judecată era să exploateze incertitudinea din sistem, ceva ce algoritmii vor face mai dificil să se întâmple. „Cu cât sentințele devin mai previzibile, cu atât rămâne mai puțin loc pentru arta pledoariei."

Când, însă, am întrebat-o pe Mandeep Dhami cum s-ar simți dacă ar fi ea însăși amenințată cu închisoarea, a dat un răspuns complet opus:

> Nu vreau ca o persoană să-și folosească intuiția atunci când ia o hotărâre în privința viitorului meu. Vreau să apeleze la o strategie bine gândită. Dorim să păstrăm independența judiciară și bunul-plac al judecătorilor de parcă ar fi ceva sacrosanct! De parcă ar fi ceva nespus de bun! Chiar dacă toate cercetările arată că nu este așa. Independența judiciară nu este deloc un lucru nemaipomenit.

Ca mai toată lumea, și eu consider că deciziile judecătorilor ar trebui să fie cât mai nepărtinitoare cu putință. Aceștia ar trebui să se orienteze după faptele săvârșite de o persoană, nu după grupul căruia se poate întâmpla să îi aparțină. Din acest punct de vedere, algoritmii nu prea sunt suficient de buni. Însă nu este de ajuns doar să indici ce este greșit la aceștia. Alegerea nu se face între un algoritm cu imperfecțiuni și un sistem imaginar perfect. Singura comparație cinstită care poate fi făcută este între un algoritm și ceea ce ne-ar rămâne în absența sa.

Cu cât m-am informat mai mult, cu cât am vorbit cu mai multe persoane, cu atât am ajuns să cred că ne așteptăm la prea mult de la judecătorii noștri, care nu sunt decât niște biete ființe din carne și oase. Inechitatea este parte

integrată a tuturor sistemelor noastre umane. Pentru fiecare Christopher Drew Brooks tratat incorect de un algoritm, există nenumărate cazuri, ca acela al lui Nicholas Robinson, în care un judecător greşeşte pe cont propriu. Dar, dacă există un algoritm – chiar şi unul imperfect – care lucrează cu judecătorii pentru a veni în sprijinul gândirii lor adesea defectuoase, acesta reprezintă, cred eu, un pas în direcţia corectă. Pentru că un algoritm bine conceput şi calibrat aşa cum trebuie poate ajuta la eliminarea tendenţiozităţii sistematice şi a erorilor aleatorii. Nu poţi schimba o întreagă cohortă de judecători, mai ales dacă nici ei înşişi nu sunt în stare să-ţi spună cum au luat deciziile pe care le-au luat.

Pentru a crea un algoritm care să fie utilizat în sistemul judiciar trebuie să ne oprim însă şi să ne gândim temeinic care este, mai exact, scopul acestui sistem. În loc să închidem ochii şi să sperăm că totul va fi bine, algoritmii ne cer să ne formăm o idee clară şi lipsită de ambiguitate despre ce vrem ca ei să realizeze negreşit, precum şi o înţelegere fermă a eşecurilor umane pe care vin să le înlocuiască. Silindu-ne, astfel, să purtăm o dezbatere dificilă despre cum ar trebui luate deciziile într-un tribunal. Nu va fi simplu, dar aceasta este cheia pentru a stabili dacă algoritmii vor putea fi vreodată suficient de buni în acest scop.

În sistemul judiciar se manifestă tensiuni care tulbură apele şi fac ca acest gen de întrebări să-şi găsească doar cu mare dificultate răspunsurile. Dar există şi alte domenii în care algoritmii pătrund treptat şi în care luarea deciziilor este mult mai puţin conflictuală, iar contribuţia obiectivă şi pozitivă pe care o pot aduce algoritmii în societate poate fi mult mai clar definită.

Capitolul 4

MEDICINA

În anul 2015, niște oameni de știință animați de spiritul pionieratului au realizat un studiu neobișnuit pentru a determina acuratețea modului în care era diagnosticat cancerul mamar.[1] Ei le-au pus la dispoziție câte un monitor cu ecran tactil (*touch-screen*) celor 16 participanți la test și le-au cerut să trieze imaginile cu țesut mamar. Probele fuseseră prelevate de la femei reale al căror țesut mamar fusese îndepărtat prin biopsie, secționat în felii subțiri și colorat cu diverse substanțe chimice pentru a evidenția vasele de sânge și canalele mamare în nuanțe de roșu, purpuriu și albastru. Participanții la test nu trebuiau decât să decidă dacă tiparele din imagini sugerau că printre celulele sănătoase se ascundeau și unele canceroase.

După o scurtă perioadă de antrenament, participanții la test au fost puși la lucru și au dat rezultate impresionante. Lucrând independent, au evaluat corect 85% dintre probe.

Atunci cercetătorii au descoperit un lucru remarcabil. Dacă începeau să cumuleze răspunsurile – combinând voturile fiecărui participant la test pentru a obține o evaluare generală a imaginii –, rata de corectitudine creștea brusc, ajungând la 99%!

Cu adevărat extraordinară la acest studiu nu era doar priceperea participanților la test. Era identitatea lor. Acești salvatori curajoși nu erau oncologi. Nu erau nici patologi. Nu erau nici asistente medicale. Nu erau nici măcar studenți la Medicină. Erau doar niște porumbei.

Locurile de muncă ale patologilor sunt încă în siguranță, cel puțin, pentru o vreme – cred că nici măcar oamenii de știință care au pus la punct studiul nu ar îndrăzni să sugereze că medicii ar trebui înlocuiți cu banalii și bine-cunoscuții porumbei. Însă experimentul a demonstrat un aspect important: descoperirea tiparelor ascunse printre grupurile de celule nu este o abilitate specific umană. Așadar, dacă un porumbel se poate descurca onorabil în rezolvarea unei asemenea sarcini, de ce n-ar putea s-o facă și un algoritm?

VÂNĂTORII DE TIPARE

Întreaga istorie a medicinei moderne și toată practicarea sa se bazează pe descoperirea de tipare în cadrul informațiilor. Încă din Grecia antică, de când Hipocrate și-a întemeiat școala de medicină, cu vreo 2 500 de ani în urmă, observarea, experimentarea și analizarea informațiilor au fost fundamentale în lupta pentru menținerea sănătății.

Înainte de aceasta, medicina – în cea mai mare parte – abia dacă se deosebea de magie. Oamenii credeau că te îmbolnăveai dacă ofensai vreun zeu sau altul și că maladia era consecința faptului că un spirit malefic îți poseda trupul. Prin urmare, activitatea unui doctor implica o mulțime de incantații, psalmodieri și respectarea unor superstiții, ceea ce sună extrem de amuzant, dar probabil că nu era deloc așa și pentru persoana care se baza pe toate acestea ca să nu-și dea duhul.

Nu vreau să spun că Hipocrate ar fi vindecat lumea, de unul singur și pentru totdeauna, de iraționalitate și superstiție (în definitiv, se zvonea, printre altele, că fiica sa ar fi fost un dragon de 30 de metri)[2], ci că a abordat medicina într-o manieră cu adevărat revoluționară. În opinia lui, cauzele bolilor trebuiau înțelese prin cercetare rațională, nu prin magie. Punând accentul pe semnalarea cazurilor și pe

observație, el a întemeiat medicina ca știință, câștigând pe merit renumele de „părinte al medicinei moderne".[3]

Cu toate că explicațiile științifice aduse de Hipocrate și de colegii săi nu prea se mai susțin dacă le așezăm sub lupa științei moderne (doar ei credeau că sănătatea nu era decât rezultatul armoniei, al echilibrului, dintre sânge, flegmă, bilă galbenă și bilă neagră!)[4], concluziile la care ajungeau analizând informațiile pe care le dețineau sunt, în mod sigur, valabile.[5] (Spre exemplu, au fost primii care au tras concluzii de genul: „Pacienții care sunt grași de la natură sunt predispuși să moară înaintea celor slabi.") Este un laitmotiv care se regăsește de-a lungul timpurilor. Cunoașterea științifică e posibil s-o fi luat-o de multe ori pe căi greșite, însă progresul se naște din capacitatea noastră de a descoperi tipare, de a clasifica simptome și de a folosi toate aceste observații pentru a prezice ce poate aduce viitorul pentru un pacient.

Istoria medicinei este plină de astfel de exemple. Spre exemplu, în China secolului al XV-lea, tămăduitorii și-au dat seama pentru prima oară că pot să imunizeze oamenii împotriva variolei. După secole de experimentare, au găsit un tipar pe care-l puteau folosi pentru a micșora de zece ori riscul de deces în urma acestei boli. Tot ce aveau de făcut era să găsească o persoană cu o formă ușoară a bolii, să-i recolteze crustele, să le usuce, să le piseze și să le sufle în nasul unei persoane sănătoase.[6] Sau să ne gândim la epoca de aur a medicinei din secolul al XIX-lea, când aceasta adopta metode din ce în ce mai științifice, iar căutarea de tipare printre informații a devenit o parte integrantă a atribuțiilor unui medic. Unul dintre ei a fost maghiarul Ignaz Semmelweis, care, în anii 1840, a remarcat ceva surprinzător în statisticile privind decesele din saloanele de maternitate. Femeile care nășteau în saloane ticsite de medici erau de cinci ori mai predispuse să se îmbolnăvească de septicemie decât cele din saloanele supravegheate de moașe. Informațiile indicau, de

asemenea, şi motivul: doctorii disecau cadavre, iar imediat după aceea se ocupau de femeile însărcinate fără să se mai oprească să se spele pe mâini![7]

Ceea ce era adevărat pentru China secolului al XV-lea şi Europa secolului al XIX-lea este adevărat şi în prezent pentru medicii din întreaga lume nu numai atunci când studiază maladiile populaţiei, ci şi când îşi îndeplinesc rolul lor principal, de zi cu zi, adică acela de a-şi îngriji pacienţii. Osul acesta este rupt sau nu? Durerea aceasta de cap este una normală sau reprezintă simptomul a ceva mai grav? Este bine să se prescrie un tratament cu antibiotice pentru a face să dispară acest abces? Toate aceste lucruri ţin de recunoaştere, clasificare şi predicţie. Aptitudini la care algoritmii se întâmplă să fie foarte, foarte buni.

Firește, sunt multe aspecte în profesiunea de medic pe care un algoritm nu va fi, probabil, niciodată în stare să le reproducă. Empatia, de exemplu. Sau abilitatea de a acorda sprijin pacienţilor atunci când se confruntă cu dificultăţi sociale, psihologice sau chiar financiare. Dar sunt şi domenii ale medicinei în care algoritmii pot să fie de ajutor. Mai ales în ipostazele în care recunoaşterea tiparelor medicale se manifestă în cea mai pură formă a sa, iar clasificarea şi predicţia sunt preţuite aproape până la excluderea a orice altceva. Mai ales într-un domeniu precum patologia.

Patologii sunt medicii pe care un pacient îi întâlneşte rareori. Ori de câte ori o probă de sânge sau de ţesut este trimisă la analiză, ei sunt cei care, stând în vreun laborator îndepărtat, o vor examina şi vor completa buletinul de analiză. Contribuţia lor vine chiar la sfârşitul procesului de diagnosticare, acolo unde priceperea, precizia şi fiabilitatea sunt de o importanţă vitală. Patologii sunt adesea cei care îţi spun dacă ai sau nu cancer. Prin urmare, dacă biopsia pe care o analizează este singurul lucru care te desparte de chimioterapie, intervenţie chirurgicală sau chiar ceva mai rău, vrei să fii sigur că ei fac lucrurile cum trebuie.

Iar munca lor nu este deloc ușoară. Potrivit sarcinilor care îi revin, un patolog oarecare va examina sute de lamele pe zi, fiecare conținând zeci de mii – uneori chiar sute de mii – de celule care stau în suspensie între plăci de sticlă micuțe. Este cel mai dificil joc de tip *Unde este Wally?** pe care ți-l poți imagina. Patologii trebuie să examineze cu meticulozitate fiecare probă în căutarea unor anomalii infime care s-ar putea ascunde oriunde în imensa galaxie de celule pe care o văd ei sub lentilele microscopului.

„Este o sarcină extraordinar de dificilă", a declarat Andy Beck[8], patolog la Universitatea Harvard și fondatorul PathAI, o companie înființată în 2016 care creează algoritmi pentru clasificarea lamelelor de biopsie. „Dacă fiecare patolog ar trebui să examineze cu foarte mare atenție cinci lamele pe zi, ai putea să crezi că ar reuși să atingă perfecțiunea. Dar nu aceasta este lumea în care trăim."

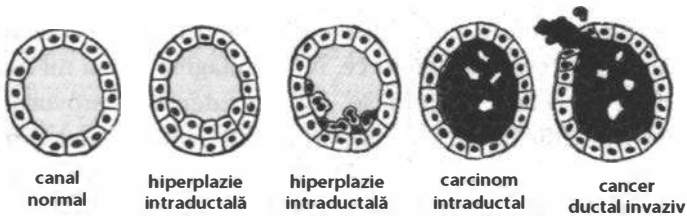
În mod cert, nu este. Iar în lumea reală, sarcina patologilor devine și mai dificilă din cauza complexității frustrante a biologiei. Să ne întoarcem la exemplul cu porumbeii și cancerul mamar pe care aceștia reușeau atât de bine să îl detecteze. A decide dacă o persoană suferă de această afecțiune nu înseamnă să dai un răspuns simplu, precum „da" sau „nu", deoarece diagnosticele de cancer mamar acoperă un întreg spectru. La unul din capete se află probele benigne, în care celulele normale arată exact așa cum ar trebui să arate. La capătul opus se află tumorile de tipul cel mai urât – carcinoamele invazive, în care celulele canceroase au migrat din canalele mamare și au început să prolifereze în țesutul înconjurător. Cazurile situate

* Colecție de cărți pentru copii creată de ilustratorul britanic Martin Handford, în care personajul numit Wally se ascunde în desenele ce înfățișează diferite mulțimi, iar cititorii trebuie să-l descopere. Editorii din diverse țări i-au dat personajului nume diferite, astfel încât acesta este cunoscut drept Waldo în SUA și Canada, drept Charlie, în Franța, Walter în Germania, Ali în Turcia, Efi în Israel și Willy în Norvegia (n. tr.).

la aceste extreme sunt relativ ușor de depistat. Un studiu recent arăta că patologii reușesc să diagnosticheze corect 96% din specimenele care sunt, în mod clar, maligne, o precizie aproximativ echivalentă cu aceea pe care a atins-o și grupul de porumbei când i s-a prezentat o sarcină similară.[9]

Între aceste două extreme, însă – între absolut normal și malign într-un mod îngrozitor de clar –, se situează alte câteva categorii, mai ambigue, așa cum se poate vedea în diagrama prezentată în continuare. În proba dumneavoastră s-ar putea să existe un grup de celule atipice care arată cam dubios, dar care nu înseamnă neapărat ceva îngrijorător. Ar putea fi niște formațiuni precanceroase care s-ar putea transforma (sau nu!) în ceva grav. Sau s-ar putea să existe un cancer care încă nu s-a răspândit dincolo de canalele mamare (așa numitul carcinom *in situ**).

Cărei categorii anume se va considera că aparține proba dumneavoastră va avea probabil un impact imens asupra tratamentului pe care îl veți urma. În funcție de locul din spectru în care se situează proba, medicul v-ar putea sugera orice, de la mastectomie până la niciun fel de intervenție.



Problema este că poate fi extrem de dificil să faci deosebirea între aceste categorii ambigue. Până și patologii cu

* *Carcinomul in situ* (CIS) este o formă timpurie de cancer, definită prin absența invaziei de celule tumorale în țesutul înconjurător. Cu alte cuvinte, celulele neoplazice proliferează în habitatul lor normal, de unde și denumirea de *in situ* (lat.), „în locul său", putând fi extirpate complet (n. tr.).

experiență pot avea opinii diferite în privința diagnosticării corecte a unei anumite probe. Pentru a verifica în ce măsură variază opiniile medicale, un studiu din 2015 a selectat 72 de biopsii ale țesutului mamar despre care se considera că ar conține celule cu anomalii benigne (o categorie situată pe la mijlocul spectrului) și apoi s-au solicitat părerile unui număr de 115 patologi. În mod îngrijorător, doar în 48% dintre cazuri patologii au emis diagnostice similare.[10]

Și, de vreme ce, pare-se, șansele de diagnosticare corectă sunt de jumi-juma, am putea la fel de bine să dăm cu banul pentru a stabili diagnosticul. Cap, și te-ai putea alege cu o mastectomie inutilă (care să te coste sute de mii de dolari dacă trăiești în SUA). Pajură, și ai putea să ratezi șansa de a prinde cancerul în stadiul incipient. Oricum ar fi, impactul poate fi devastator.

Evident, când miza este atât de mare, cel mai mult contează acuratețea. Așadar, un algoritm poate să obțină rezultate mai bune decât acestea?

MAȘINI CARE VĂD

Până de curând, construirea unui algoritm care să poată recunoaște ceva, indiferent ce, într-o imagine, ca să nu mai vorbim de celule canceroase, era considerată o provocare extrem de dificilă. Nu avea importanță că oamenii înțelegeau fotografiile cu atâta ușurință; explicarea exactă a *modului în care reușeau* să facă acest lucru s-a dovedit o sarcină neînchipuit de dificilă.

Ca să înțelegeți de ce, imaginați-vă că scrieți instrucțiuni pentru a-i comunica unui computer dacă într-o fotografie apare sau nu un câine. Ați putea începe cu ceea ce este evident: dacă are patru picioare, urechi pleoștite, blană și așa mai departe. Dar ce ar trebui făcut cu acele fotografii în care câinele stă jos? Sau cu acelea în care nu i se văd toate picioarele? Dar cu pozele care înfățișează un câine cu urechi

ascuțite? Sau unul cu urechile pleoștite doar pe jumătate? Sau altul care nu privește deloc spre obiectivul aparatului de fotografiat? Și prin ce diferă „blana" unui câine dintr-o fotografie de imaginea unui covor cu fire lung? Sau de lâna unei oi? Sau de iarbă?

Sigur că ați putea introduce toate aceste detalii ca instrucțiuni suplimentare și să faceți apoi algoritmul să treacă prin fiecare tip posibil de urechi de câine, blană de câine sau poziție șezândă, numai că, în scurt timp, algoritmul va deveni atât de voluminos, încât va fi complet nefuncțional înainte chiar ca dumneavoastră să fi început măcar să faceți deosebirea între câini și alte creaturi cu blană și patru picioare. Așadar, trebuie să găsiți o altă modalitate de diferențiere. Secretul este să renunțați la o paradigmă bazată pe reguli și să folosiți ceva ce se numește o „rețea neurală".[11]

Vă puteți imagina rețeaua neurală ca pe o imensă structură matematică în care apar o mulțime de butoane și cadrane. La un capăt se introduce poza, aceasta parcurge întreaga structură, iar la capătul celălalt iese o presupunere despre ce ar putea conține imaginea respectivă. O probabilitate pentru fiecare categorie: Câine sau Non-Câine.

La început, rețeaua neurală va fi o mare harababură. Începe prin a nu ști nimic – nu are habar ce este sau ce nu este un câine. Toate butoanele și cadranele sunt setate în poziția „aleatoriu". În consecință, răspunsurile pe care le va oferi vor fi complet aiurea – rețeaua nu ar putea recunoaște o imagine nici dacă însăși sursa ei de alimentare cu electricitate ar depinde de asta. Însă, cu fiecare fotografie pe care o veți introduce, veți ajusta butoanele și cadranele acelea. Treptat, o veți antrena.

Și acum puteți să introduceți imaginea unui câine. După fiecare supoziție pe care o va face rețeaua, un set de reguli matematice va intra în funcțiune pentru a ajusta toate butoanele, până când predicția se va apropia de răspunsul corect. Atunci veți introduce o altă imagine și apoi încă

una, făcând ajustări de fiecare dată când rețeaua neurală va înțelege ceva greșit, întărind căile prin seria de butoane care conduc la succes și reducând, treptat, importanța acelora care au condus la eșec. Informația despre ceea ce face ca o poză reprezentând un câine să semene cu o altă poză înfățișând un câine se propagă în sens invers prin rețea. Procesul continuă până când – după introducerea a sute și mii de fotografii – se primesc tot mai puține răspunsuri greșite cu putință. În cele din urmă, îi puteți arăta rețelei o imagine pe care nu a văzut-o niciodată și aceasta va reuși să afirme, cu un nivel ridicat de acuratețe, dacă aceasta înfățișează sau nu un câine.

Surprinzător în ceea ce privește rețelele neurale este faptul că utilizatorii acestora, de obicei, nu înțeleg cum sau de ce ajunge algoritmul la concluziile sale. O rețea neurală care clasifică imagini de câini nu funcționează prin identificarea trăsăturilor pe care dumneavoastră sau eu le-am putea recunoaște ca fiind potențial canine. Ea nu caută o unitate de măsură pentru „*chihuahua*-ism" sau „marele danez-ism" –, totul este mult mai abstract decât atât: rețeaua neurală detectează tipare ale marginilor și ale luminii și umbrei din fotografii, ceea ce nu este foarte inteligibil pentru un observator uman (aruncați o privire asupra exemplului de recunoaștere a imaginii din capitolul „Puterea" ca să înțelegeți mai bine ce vreau să spun). Pentru că unei ființe umane îi este greu să conceptualizeze acest proces, programatorii nu știu decât că au setat algoritmul să dea răspunsurile corecte; ei nu știu neapărat în amănunt cum ajunge acesta la rezultatul respectiv.

Iată încă un „algoritm de învățare automată", ca acele „păduri aleatorii" (*random forests*) pe care le-am întâlnit la capitolul „Justiția". El merge dincolo de ceea ce inginerul îl programează să execute și învață singur din imaginile care i se furnizează. Această capacitate de învățare înzestrează algoritmul cu o remarcabilă „inteligență artificială". Iar

multiplele niveluri de butoane și cadrane îi conferă rețelei o structură în profunzime, de unde și sintagma *deep learning**.

Rețelele neurale există de pe la mijlocul secolului al XX-lea, însă până de curând ne-a lipsit accesul pe scară largă la calculatoarele cu adevărat puternice necesare pentru a obține tot ce este mai bun de la aceste rețele. Lumea a fost, în cele din urmă, nevoită să își manifeste poziția și să le ia în serios în 2012, când cercetătorul în calculatoare Geoffrey Hinton și doi studenți de-ai săi au înscris un nou tip de rețea neurală într-un concurs de recunoaștere a fotografiilor.[12] Provocarea care li se lansase – printre altele – era recunoașterea câinilor. Algoritmul lor înzestrat cu inteligență artificială i-a făcut praf până și pe cei mai buni concurenți și a declanșat un reviriment de mari proporții în domeniul *deep learning*.

Un algoritm care lucrează fără ca noi să știm cum ia deciziile s-ar putea să semene a scamatorie, dar poate că acest lucru nu este chiar atât de diferit de modul în care învățăm noi înșine. Gândiți-vă la comparația următoare. O echipă a instruit recent un algoritm să facă deosebirea între fotografiile cu lupi și acelea în care apar câini husky. Apoi au arătat cum, datorită modului în care își reglase propriile cadrane, algoritmul nu folosea absolut deloc indicii care să aibă legătură cu vreun câine. Își baza răspunsul pe informația dacă imaginea avea sau nu zăpadă în fundal. Plus zăpadă = lup. Fără zăpadă = husky.[13]

La puțin timp după publicarea lucrării lor, stăteam de vorbă cu Frank Kelly, un profesor de matematică de la Universitatea Cambridge, care mi-a relatat o conversație pe care o avusese cu nepotul său. Își ducea nepotul de patru ani la grădiniță, când au trecut pe lângă un câine husky. Nepotul său a remarcat că acel câine „arăta" ca un lup. Când Frank l-a întrebat de unde știa că nu era un lup, micuțul îi răspunsese: „Pentru că e în lesă."

* „Învățare profundă" (engl.) (vezi glosarul) (n. tr.).

Sunt două lucruri de dorit de la un algoritm performant de scanare a cancerului mamar. În primul rând, să fie suficient de *sensibil* ca să detecteze anomaliile prezente la toți sânii care prezintă tumori, fără să-i scape vreun pixel din imagine și fără să anunțe apoi că nu există anomalii. Dar, în al doilea rând, mai trebuie și ca algoritmul să fie destul de *precis*, să dea dovadă de *specificitate*, astfel încât să nu semnaleze că o porțiune de țesut mamar perfect normal ar fi demn să stârnească suspiciuni.

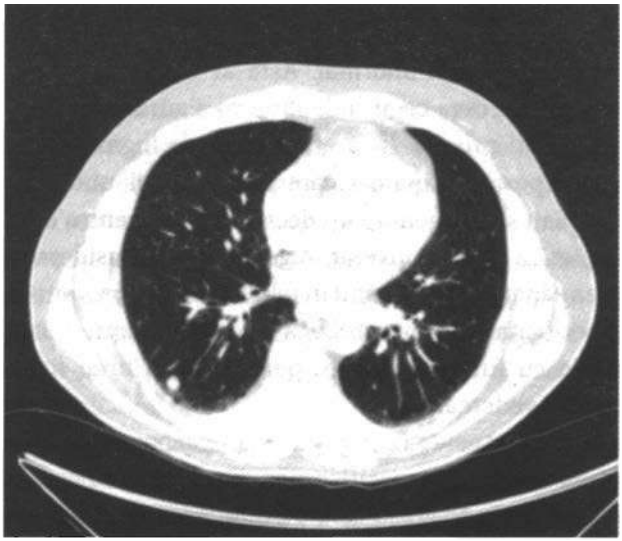
Am întâlnit deja principiile sensibilității și specificității în capitolul „Justiția". Sunt rude apropiate cu *fals negativele* și *fals pozitivele* (sau cu Darth Vader-ii și a Luke Skywalker-ii – denumirile care, dacă mă întrebați pe mine, ar trebui să fie folosite, în mod oficial, și în literatura științifică). În contextul nostru, al diagnosticării medicale, un fals pozitiv apare atunci când unei femei sănătoase i se spune că are cancer la sân, iar un fals negativ – atunci când o femeie cu tumori primește diagnosticul că ar fi sănătoasă. Un test specific aproape că nu va avea rezultate fals pozitive, pe când unul sensibil va avea puține rezultate fals negative. Nu are importanță în ce context lucrează algoritmul – dacă prognozează că un anumit infractor va recidiva sau nu, dacă diagnostichează cancerul mamar sau (cum vom vedea în capitolul „Infracționalitatea") dacă identifică tiparele activității infracționale – povestea este mereu aceeași. E de dorit să avem cât mai puține rezultate fals pozitive și fals negative cu putință.

Problema este că perfecționarea unui algoritm înseamnă adesea o alegere între sensibilitate și specificitate. Dacă vă concentrați asupra îmbunătățirii uneia, asta înseamnă adesea o scădere a performanțelor celeilalte. Dacă, de exemplu, ați decis să dați întâietate eliminării complete a fals negativelor,

algoritmul ar putea să semnaleze fiecare țesut mamar pe care îl vede drept anormal. Asta ar însemna că are 100% sensibilitate, ceea ce ar îndeplini, cu siguranță, obiectivul pe care vi l-ați propus. Însă ar însemna și îngrozitor de mulți oameni perfect sănătoși care ar fi supuși unui tratament inutil. Sau să spunem că ați decis să optați pentru eliminarea completă a fals pozitivelor. Algoritmul ar considera pe toată lumea sănătoasă, atingând în felul acesta un procent de 100% la specificitate. Minunat! Dacă nu cumva sunteți una dintre femeile cu tumori mamare pe care algoritmul tocmai le-a trecut cu vederea.

Interesant este că patologii umani nu par să aibă probleme în ceea ce privește specificitatea. Ei aproape că nu identifică niciodată în mod greșit celule sănătoase ca fiind canceroase. Însă oamenii au o oarecare problemă cu sensibilitatea: pur și simplu, se pare că ne este îngrijorător de ușor să trecem cu vederea micile tumori, ba chiar și pe cele care sunt, în mod evident, maligne.

Aceste slăbiciuni omenești au fost scoase la iveală într-o competiție desfășurată recent în care ființele umane s-au confruntat cu algoritmii. Echipe de programatori din toată lumea s-au luat la întrecere cu patologi, să găsească toate tumorile existente pe patru sute de lamele, într-un concurs cunoscut sub numele de CAMELYON16. Ca să ușureze lucrurile, toate cazurile se situau la una dintre cele două extreme: țesut perfect normal sau cancer mamar invaziv, iar patologii nu aveau nicio limită temporală: dispuneau de tot timpul de care aveau nevoie pentru a examina biopsiile. Așa cum era de așteptat, oamenii au pus diagnosticul general în mod corect (cu o precizie de 96%)[14] și fără să identifice niciun fals pozitiv. Însă le-au scăpat și o mulțime de mici celule canceroase ascunse prin țesut, reușind să observe numai 73% dintre acestea în 30 de ore de examinare.



Problema nu a fost neapărat numărul efectiv de pixeli care trebuiau verificați. Oamenii pot cu ușurință să nu observe anomalii foarte evidente chiar și atunci când se uită direct la ele. În anul 2013, cercetătorii de la Harvard au ascuns în secret imaginea unei gorile într-o serie de radiografii toracice și au cerut unui număr de 24 de radiologi, care nu bănuiau nimic, să verifice radiogramele respective pentru a depista semne de cancer. Ei bine, 83% dintre aceștia nu au observat gorila, în pofida faptului că sistemul de urmărire a privirii folosit de oamenii de știință a arătat că majoritatea se uitaseră chiar la ea.[15] Puteți încerca și dumneavoastră să o depistați, în poza de mai sus.[16]

Algoritmii au exact problema opusă: vor marca imediat grupurile de celule cu anomalii, ba chiar și pe unele perfect sănătoase. În timpul studiului CAMELYON16, de exemplu, cea mai performantă rețea neurală participantă a reușit să identifice un procent impresionant (92,4%) dintre tumori[17], însă a făcut astfel opt erori de identificare fals pozitivă per lamelă, prin semnalarea greșită a unor grupuri de celule

normale ca fiind canceroase. Cu o specificitate atât de scăzută, algoritmii actuali care reprezintă cel mai avansat stadiu al cercetării înclină, în mod cert, în diagnosticare, spre abordarea „toată lumea are cancer mamar!", ceea ce face, pentru moment, ca ei să nu fie, practic, suficient de performanți pentru a putea să emită propriile rapoarte patologice.

Veștile bune sunt, însă, că nici nu ne propunem să le cerem acest lucru! Intenționăm să combinăm, în schimb, calitățile omului cu acelea ale mașinii. Algoritmul efectuează munca de rutină – examinarea volumului imens de informații de pe lamele – evidențiind câteva arii principale de interes. Apoi patologul preia comanda. Nu are importanță dacă mașina semnalează celule care nu sunt canceroase, pentru că specialistul uman va face o verificare rapidă, eliminând tot ce este normal. Genul acesta de parteneriat în care algoritmul realizează o pre-examinare, o pre-sortare a informațiilor, nu numai că economisește foarte mult timp, ci crește și precizia generală a diagnosticului până la un procent uimitor de 99,5%.[18]

Oricât de miraculos vi s-ar părea, adevărul este că patologii umani au fost întotdeauna buni la diagnosticarea tumorilor canceroase agresive. Cazurile dificile pentru ei sunt cele ambigue, de la mijlocul spectrului, în care deosebirea dintre celulele canceroase și cele sănătoase este mai subtilă. Pot, oare, algoritmii să ne fie de ajutor și în acest caz? Răspunsul este (probabil) da. Însă nu încercând să stabilească un diagnostic prin utilizarea categoriilor complicate pe care patologii le-au folosit dintotdeauna. Poate că algoritmul – care este cu mult mai performant în depistarea anomaliilor ascunse în mici fragmente de informație – ar putea să ofere, în schimb, o cu totul altă modalitate, mai bună, de diagnosticare. Făcând ceva ce medicii umani nu pot.

STUDIUL REALIZAT PE CĂLUGĂRIȚE*

În anul 1986, David Snowdon, un epidemiolog de la Universitatea din Kentucky, a reușit să convingă 678 de călugărițe să-i cedeze creierele după moarte. Măicuțele, care aparțineau Ordinului Surorilor de Notre Dame, au acceptat să participe la o cercetare științifică ieșită din comun despre cauzele apariției maladiei Alzheimer.

Fiecare dintre aceste femei, care la începutul studiului aveau vârste cuprinse între 75 și 103 ani, avea să efectueze o serie de teste de memorie în fiecare an pe durata întregii vieți. Apoi, după moarte, creierele lor aveau să fie donate proiectului respectiv. Indiferent dacă manifestau sau nu simptome de demență, ele au promis că vor permite echipei lui Snowdon să înlăture cel mai prețios organ al lor pentru a fi analizat în căutarea semnelor acestei afecțiuni.[19]

Generozitatea călugărițelor a dus la crearea unui remarcabil set de date. Din moment ce niciuna dintre ele nu avea copii, nu fuma și nu bea prea mult, oamenii de știință au putut să excludă mulți dintre factorii externi despre care se credea că ar fi dus la creșterea probabilității de a contracta maladia Alzheimer. Și, din moment ce aveau un stil de viață asemănător, cu acces egal la sistemul de sănătate și la cel de asistență socială, călugărițele ofereau ele însele și grupul de control.

Totul a decurs bine când, după câțiva ani de la începerea studiului, echipa a descoperit că acest grup experimental mai oferea un tezaur de informații la care putea avea acces. În tinerețe, multora dintre măicuțele acum în etate li se ceruse, înainte de a li se permite să se călugărească, să înmâneze Ordinului un eseu autobiografic scris de mână. Aceste eseuri fuseseră scrise când femeile aveau – în medie – doar 22 de

* *Nun Study of Aging and Alzheimer's Disease* (engl.), „Studiul privind îmbătrânirea și [instalarea maladiei lui] Alzheimer realizat pe călugărițe", demarat în 1986 de David Snowdon (n. red.).

ani, cu zeci de ani înainte ca oricare dintre ele să manifeste vreun simptom de demență. Cu toate acestea, în mod uimitor, oamenii de știință au descoperit în scrisul lor de mână indicii care preziceau ce avea să li se întâmple într-un viitor îndepărtat.

Cercetătorii au analizat complexitatea limbajului din fiecare eseu și au găsit o legătură între cât de elocvente erau călugărițele ca tinere femei și șansele ca ele să dezvolte demență la o vârstă înaintată.

Iată, de exemplu, un citat dintr-un eseu al unei călugărițe care și-a păstrat abilitățile cognitive excelente întreaga sa viață:

> După ce am terminat clasa a VIII-a, în 1921, am dorit să devin aspirantă la Mankato, însă nu am avut curajul să cer eu însămi permisiunea părinților, așa încât sora Agreda a făcut-o în locul meu, iar aceștia și-au dat imediat consimțământul.

Comparați această frază cu următoarea, scrisă de o altă călugăriță, ale cărei punctaje de acuitate a memoriei au decăzut constant în ultimii ani de viață:

> După ce am absolvit școala, am lucrat la oficiul de poștă.

Asocierea era atât de clară, încât oamenii de știință puteau de-acum să prezică de-a dreptul care dintre călugărițe ar fi putut avea demență numai citindu-le scrisorile. 90% dintre cele care au dezvoltat mai târziu maladia Alzheimer aveau o „capacitate lingvistică limitată" ca tinere femei, pe când numai 13% dintre călugărițele care și-au păstrat capacitățile lingvistice până la o vârstă înaintată au primit un punctaj de „densitate scăzută a ideilor" la eseurile lor.[20]

Unul dintre lucrurile pe care acest studiu le pune în lumină este cantitatea incredibilă de informații pe care încă o avem de aflat despre corpurile noastre. Chiar dacă știm că acea legătură ar putea să existe, nu știm și *de ce*. (Oare o bună educație este suficientă pentru a împiedica apariția demenței? Sau persoanele predispuse să se îmbolnăvească de Alzheimer

se simt mai în largul lor folosind un limbaj mai simplu?) Dar asta ar sugera că maladia Alzheimer se dezvoltă pe o durată de câțiva zeci de ani!

Încă și mai important pentru ceea ce ne propunem, asta demonstrează că indicii subtile despre viitoarea noastră stare de sănătate se pot ascunde chiar și în cele mai mărunte, mai neașteptate fragmente de informație, cu ani și ani înainte ca noi să manifestăm simptomele vreunei maladii. Aceasta ne oferă o idee despre cât de puternici ar putea fi algoritmii medicali care vor fi folosiți în viitor pentru analizarea informațiilor. Poate că, într-o bună zi, vor reuși să detecteze semnele cancerului cu câțiva ani înaintea medicilor.

PUTEREA PREDICȚIEI

La sfârșitul anilor 1970, un grup de patologi din Ribe, un oraș din Danemarca, a început să efectueze duble mastectomii pe un lot de cadavre. Femeile decedate aveau vârste cuprinse între 22 și 89 de ani, iar șase din totalul de 83 muriseră de cancer mamar invaziv. Bineînțeles că atunci când cercetătorii au pregătit sânii extirpați în vederea examinării lor de către un patolog – împărțind fiecare sân în patru bucăți, apoi tăind țesutul în felii subțiri pentru a-l așeza pe lamele –, cele șase eșantioane au manifestat semnele maladiei. Dar, spre uimirea cercetătorilor, din restul de 77 de femei – care muriseră din alte cauze, fără nicio legătură cu cancerul, printre care boli de inimă și accidente de automobil –, aproape un sfert prezentau semnele de avertizare ale cancerului mamar pe care patologii le căutau la pacientele în viață.

Fără să manifeste niciun semn de boală, 14 femei aveau celule canceroase *in situ* care nu se dezvoltaseră niciodată dincolo de canalele sau glandele mamare. Celule care ar fi fost considerate drept cancer mamar malign dacă femeile ar fi fost

în viață. Trei aveau celule atipice care ar fi fost, de asemenea, considerate a reprezenta un motiv de îngrijorare la o biopsie, iar o femeie suferea chiar de cancer mamar invaziv, deși nu avusese nici cea mai vagă idee despre acest lucru atunci când murise.[21]

Cifrele au fost surprinzătoare, dar studiul nu era doar o coincidență. Și alți cercetători găsiseră rezultate asemănătoare. De fapt, unii chiar estimau că, în orice moment, în jur de 9% dintre femei ar putea să trăiască fără să știe că au dezvoltat tumori mamare[22] – cam de zece ori mai mult față de cele care sunt diagnosticate în prezent cu cancer mamar.[23]

Așadar, ce se întâmplă? Avem de-a face cu o epidemie tăcută? Conform doctorului Jonathan Kanevsky, un inovator în medicină și chirurg rezident la Universitatea McGill din Montreal, răspunsul este: „Nu. Nu chiar." Pentru că prezența cancerului nu este neapărat o problemă:

> Dacă are cineva o celulă canceroasă în corp, sunt șanse ca sistemul imunitar să o identifice drept o celulă mutantă, să o atace și să o distrugă, iar acel cancer nu se va dezvolta niciodată, devenind ceva grav. Uneori, însă, sistemul imunitar se zăpăcește, adică organismul favorizează dezvoltarea cancerului, permițându-i să prolifereze. În acest moment, cancerul poate să ucidă.[24]

Nu toate tumorile sunt la fel. De unele se va ocupa organismul, altele vor rămâne liniștite acolo până la trecerea în neființă, iar altele ar putea să devină un cancer agresiv în toată regula. Problema este că, adesea, avem foarte puține căi de a ști ce se va întâmpla în fiecare caz în parte.

Este și motivul pentru care aceste categorii dificile, situate între benign și malign pot fi o problemă. Sunt singurele clasificări pe care medicii le au la dispoziție, dar, dacă unul dintre ei găsește în biopsia dumneavoastră un grup de celule care îi dau, cât de cât, de bănuit, eticheta pe care va alege să i-o pună poate fi utilă doar pentru a descrie ce există în țesut în prezent. Asta nu este de prea mare ajutor atunci când trebuie

să-ți furnizeze indicii despre viitor. Și firește că viitorul îi preocupă cel mai mult pe pacienții îngrijorați.

În consecință, oamenii sunt, adesea, extrem de precauți în privința tratamentului pe care îl aleg. Să luăm, de exemplu, cancerele *in situ*. Această categorie este plasată spre capătul mai îngrijorător al spectrului, căci deja este prezentă o excrescență canceroasă care încă nu s-a răspândit în țesutul înconjurător. Oricât de grav ar suna prognosticul de „cancer *in situ*", numai aproximativ unul din zece se va transforma în ceva care ar putea să ucidă. Nu este mai puțin adevărat că, în Statele Unite, un sfert dintre femeile care primesc acest diagnostic vor fi supuse unei mastectomii totale – o intervenție majoră care le schimbă viața nu doar în plan fizic, ci, adesea, și în cel emoțional.[25]

Din ceea ce s-a constatat, cu cât faci examinări mai insistente pentru depistarea cancerului mamar, cu atât afectezi mai multe femei care și-ar fi trăit mai departe fericite viața fără să aibă habar de existența unor tumori inofensive. Un grup independent din Marea Britanie a ajuns la concluzia că, pentru fiecare 10 000 de femei care vor fi chemate să-și facă un examen mamografic în următorii 20 de ani, vor fi prevenite 43 de decese de cancer mamar. Iar un studiu publicat în *New England Journal of Medicine* a ajuns la concluzia că, pentru fiecare 10 000 de femei care efectuează mamografii de rutină, unui număr de doar 30 le vor fi depistate tumori care ar putea să le pună în pericol viața.[26] Numai că – depinzând de statisticile la care te raportezi – de trei sau patru ori mai multe femei vor fi supra-diagnosticate, fiind tratate pentru tumori care nu ar fi devenit niciodată letale.[27]

Problema supra-diagnosticării și a tratării excesive este dificil de rezolvat atunci când te pricepi la depistarea anomaliilor, dar nu și la prezicerea modului în care vor evolua ele. Chiar și așa, ar putea exista speranțe. Poate că – la fel ca în cazul călugărițelor care scriseseră eseuri în tinerețe – mici indicii despre modul în care va evolua sănătatea unei persoane

în viitor, peste ceva ani, se pot ascunde și pot fi depistate în informații trecute și prezente despre aceasta. Astfel, extragerea acestor informații ar fi sarcina perfectă pentru o rețea neurală.

Într-un domeniu în care medicii s-au chinuit zeci de ani să descopere *de ce* o anumită anomalie este mai periculoasă decât alta, un algoritm care să nu fie instruit să caute *ceva anume* ar putea să-și arate adevărata valoare. Atât timp cât poți strânge un număr foarte mare de lamele de biopsie (care să includă câteva probe de tumori care au format metastaze – adică s-au extins la alte organe din corp – și, de asemenea, câteva care nu au format metastaze) pentru antrenarea rețelei tale neurale, aceasta ar putea căuta indicii ascunse cu privire la starea de sănătate viitoare, complet liberă de oricare dintre prejudecățile care vin la pachet cu teoria. Jonathan Kanevsky considera:

> Ar depinde de algoritm să stabilească caracteristicile particulare din fiecare imagine care prezic dacă o tumoare va evolua sau nu spre metastază.[28]

Dacă ar exista un astfel de algoritm, categoria în care se încadrează biopsia dumneavoastră ar deveni mult mai puțin importantă. Nu v-ar mai preocupa „de ce" și „cum", ați putea sări direct la informația care contează: aveți sau nu nevoie de tratament?

Vestea cea bună este că munca pentru crearea unui asemenea algoritm a început deja. Andy Beck, patologul de la Harvard și președintele companiei PathAI, pe care am menționat-o mai devreme, și-a instruit algoritmul să funcționeze pe o serie de probe ale unor pacienți din Olanda și a descoperit că, de fapt, cei mai buni indicatori capabili să prezică șansele de supraviețuire nu se găseau în cancerul ca atare, ci în anomaliile din țesuturile învecinate zonei canceroase.[29] Este o descoperire importantă – un exemplu concret al modului în care algoritmii înșiși duc cercetarea mai departe, dovedind că pot să descopere tipare care să ne îmbunătățească puterea de predicție.

Dispunem, în prezent, de o bogăție incredibilă de informații pe care le putem valorifica. Datorită mamografiilor de rutină care se efectuează în întreaga lume avem, probabil, mai multe imagini ale țesutului mamar decât ale oricărui alt organ din corp. Eu nu sunt patolog, însă fiecare specialist cu care am discutat m-a convins că suntem foarte aproape de a reuși să prezicem cu toată încrederea dacă un eșantion problematic va deveni sau nu canceros. Sunt chiar șanse mari ca, până la lansarea în librării a acestei cărți, cineva, undeva, să fi transformat deja această idee revoluționară în realitate.

DIAGNOSTICAREA DIGITALĂ

Aceste idei au aplicații care depășesc diagnosticarea cancerului mamar. Rețelelor neurale pe care le construiesc Andy Beck și alții nu le pasă, în mod deosebit, ce anume examinează: câini, pălării sau bucăți de brânză. Atât timp cât le anunți când au dreptate și când greșesc, ele vor învăța. Iar acum, când această familie de algoritmi este suficient de performantă pentru a fi utilizată, ea are impact într-o mulțime de domenii ale medicinei moderne.

Un succes deosebit de important a dobândit de curând echipa Google Brain, de exemplu, care a construit un algoritm capabil să depisteze principala cauză a orbirii la nivel mondial, retinopatia diabetică, o afecțiune care poate fi prevenită. Această maladie afectează vasele de sânge din zonele sensibile la lumină ale ochiului. Dacă știi că suferi de așa ceva, ți se pot prescrie injecții pentru salvarea vederii, însă, dacă nu este depistată la timp, retinopatia diabetică poate duce la orbire ireversibilă. În India, unde accesul la specialiști care să diagnosticheze această afecțiune este limitat, 45% dintre persoanele cu retinopatie diabetică își vor pierde, într-o anumită măsură, vederea chiar înainte de a afla că suferă de așa ceva. Algoritmul echipei Google, care a fost

construit în colaborare cu medici din India, este în prezent tot atât de bun pentru diagnosticarea acestei afecțiuni ca un medic oftalmolog.

În mod similar, există algoritmi care depistează afecțiunile cardiovasculare[30], emfizemul pulmonar[31], accidentele vasculare cerebrale[32] și melanoamele pielii[33]. Există chiar și sisteme care diagnostichează în timp real* polipii în timpul unei colonoscopii pe viu.

Adevărul este că, dacă poți să fotografiezi ceva și să-i pui o etichetă, atunci poți crea și un algoritm care să descopere acel lucru. Și este posibil să obții un diagnostic mai precis (și care să fie depistat, poate, în fazele incipiente ale bolii) decât ar putea s-o facă orice medic uman.

Ce se întâmplă, însă, cu informațiile medicale care apar într-o formă mai dezordonată? Succesul acestor algoritmi de diagnosticare poate fi dus mai departe, dincolo de acest gen de sarcini specializate și concentrate asupra unor sarcini restrânse? Ar putea, oare, o mașină să atribuie un înțeles notițelor mâzgălite de un medic, de exemplu? Sau să sesizeze indiciile mărunte din felul în care un pacient își descrie durerea pe care o resimte?

Ce spuneți, însă, de cea mai nouă fantasmagorie științifico-fantastică din domeniul sănătății, în care o mașină aflată în cabinetul medicului dumneavoastră de familie ascultă cu atenție simptomele pe care le acuzați și vă analizează antecedentele medicale? Să îndrăznim să ne imaginăm un computer care stăpânește fiecare fragment de cercetare medicală de ultimă oră? Unul care oferă un diagnostic exact și un plan de tratament perfect adaptat?

Pe scurt: ce-ați spune de ceva care să semene cu Watson de la IBM?

* *Real time* (engl.), atunci când este vorba despre un sistem de prelucrare a datelor, această expresie semnifică faptul că rezultatele sunt obținute concomitent cu evenimentul care a generat datele de intrare (inputul) (n. red.).

În anul 2004, Charles Lickel se înfrupta dintr-o friptură împreună cu alți colegi de-ai săi la un restaurant din New York. Pe la jumătatea mesei, zona aceea a localului a început să se golească. Intrigat, Charles a urmărit mesenii și a descoperit că se înghesuiseră în jurul unui televizor, urmărind cu interes popularul concurs american *Jeopardy*. Celebrul câștigător al concursului, Ken Jennings, avea șansa să continue seria celor șase luni de victorie neîntreruptă care depășea toate recordurile, iar clienții localului nu doreau să rateze această performanță.[34]

Charles Lickel era vicepreședintele responsabil de software al IBM. În ultimii câțiva ani, de când Deep Blue îl învinsese pe Garry Kasparov la șah, șefii de la IBM îl pisau pe Charles să găsească o nouă provocare demnă de atenția companiei. Pe când se afla în restaurantul din New York, urmărindu-i pe clienții captivați de acel campion uman al concursului *Jeopardy*, Lickel a început să se întrebe dacă nu putea fi construit un calculator care să-l învingă.

Nu avea să fie ușor. Construirea mașinii cunoscute drept „Watson", pe care și-a imaginat-o Charles în acel restaurant, a durat șapte ani. Dar, în cele din urmă, Watson avea să-l provoace la o competiție pe Ken Jennings, într-un episod special din *Jeopardy* și să-l învingă fără echivoc la jocul care îi adusese notorietatea. Între timp, IBM a început construirea primului calculator de diagnosticare atotștiutor din lume, la care ne vom întoarce în curând. Dar, mai întâi, permiteți-mi să vă prezint câteva aspecte esențiale despre mașina care a câștigat concursul *Jeopardy*, care au stat și la baza algoritmilor de diagnosticare medicală.

Pentru cei care nu au auzit niciodată de ea, *Jeopardy* este o foarte cunoscută emisiune de divertisment americană sub forma unui concurs de întrebări inversat: concurenților li se oferă indicii sub formă de răspunsuri, iar ei trebuie

să formuleze întrebarea care a stat la baza acestora. De exemplu, la categoria „cuvinte cu sens contrar", un indiciu ar putea fi ceva de genul:

> Un element de fixare pentru a prinde ceva sau pentru a îndoi, a deforma și a elibera brusc ceva, prin intermediul căldurii sau al presiunii.

Concurentul-algoritm ar trebui să învețe să treacă prin câteva niveluri pentru a obține răspunsul corect, respectiv: „Ce este o «chingă»?" Mai întâi, Watson ar trebui să înțeleagă suficient de bine limbajul pentru a desprinde înțelesul din întrebare și a-și da seama că „element de fixare", „a prinde", „a îndoi", „a deforma" și „a elibera brusc" sunt elemente separate care compun indiciul. Aceasta, în sine, este o provocare uriașă pentru un algoritm.

Și acela nu reprezenta decât primul pas! După aceea, Watson trebuia să caute potențiali candidați care să se potrivească fiecărui indiciu în parte. „Element de fixare" ar putea evoca tot felul de posibile răspunsuri: „clemă", „buton", „agrafă" și „legătură", dar și „chingă", de exemplu. Watson trebuia să analizeze pe rând fiecare posibilitate și să calculeze cum se potrivea aceasta cu restul indiciilor. Așadar, cu toate că este puțin probabil să găsești dovezi că „agrafa" s-ar putea asocia cu indicii precum „a îndoi" și „a deforma", cuvântul „chingă" sigur se asociază cu ele, ceea ce crește siguranța lui Watson că acesta este un posibil răspuns. În final, odată ce toate elementele au fost combinate, Watson trebuia să-și susțină părerea cu fapte, alegând un singur răspuns.

Provocarea de a juca *Jeopardy* este mai banală decât aceea de a diagnostica o afecțiune, însă ambele necesită câteva mecanisme logice comune. Imaginați-vă că vă duceți la doctor, acuzând o scădere în greutate neintenționată, dureri de stomac și (ca să fie tacâmul complet!) ceva arsuri în capul pieptului. Ca și atunci când joci *Jeopardy*, provocarea o constituie să găsești posibile diagnostice (răspunsuri) care ar putea explica

simptomele (indiciile), să cauți dovezi suplimentare pentru fiecare în parte și să-ți reiterezi încrederea într-un anumit răspuns pe măsură ce devin disponibile mai multe informații. Doctorii numesc asta „diagnostic diferențial". Matematicienii o numesc „inferență bayesiană"*.

Chiar și după crearea încununată de succes a lui Watson, câștigătorul concursului de cultură generală, reconfigurarea sa ca geniu medical nu a fost o sarcină ușoară. Când, însă, IBM și-a declarat public intenția de a pătrunde în domeniul sănătății, nu s-a sfiit să facă promisiuni măreție. A anunțat întreaga lume că misiunea supremă a lui Watson era „eradicarea cancerului"[35] și l-a angajat pe celebrul actor John Hamm să îl laude, numindu-l „una dintre cele mai puternice unelte create de specia umană".

Aceasta este, cu siguranță, viziunea unei utopii medicale menite să ne inspire pe toți. Cu excepția faptului că – așa cum probabil deja știți – Watson nu s-a prea ridicat la nivelul așteptărilor.

Mai întâi, un contract prestigios încheiat cu Centrul pentru studierea cancerului „M.D. Anderson" al Universității din Texas a fost reziliat în 2016. Se zvonește că nici după ce au plătit 62 de milioane de dolari pe tehnologie[36] și au petrecut patru ani lucrând la el, Watson nu a reușit să facă mai mult decât teste de verificare strict controlate. Apoi, la sfârșitul lui septembrie 2017, o investigație a site-ului de știri din domeniul sănătății STAT a descoperit că Watson „încă se chinuia cu etapa de bază a învățării despre diferitele forme de cancer".[37]

Îngrozitor, ce mai!

Cinstit vorbind, nu au fost numai vești proaste. În Japonia, Watson reușise să diagnosticheze o femeie care avea o formă rară de leucemie, atunci când medicii nu izbutiseră acest lucru.[38] Iar analizele făcute de Watson au dus la descoperirea

* Mai multe despre Bayes veți afla în capitolul „Automobilele autonome" (n. a.).

a cinci gene legate de boala neuronilor motorii, denumită și scleroza laterală amiotrofică.[39]

În linii mari, însă, programatorii de la IBM nu au prea reușit să respecte promisiunile departamentului lor de marketing, care fusese atât de entuziast.

Este greu să nu simți o oarecare simpatie față de oricine încearcă să construiască o astfel de mașină. Teoretic, este posibil să concepi una care poate să diagnosticheze diverse afecțiuni (și chiar să le ofere pacienților planuri de tratament de bun-simț), acesta fiind un scop admirabil în sine. Numai că este și foarte greu. Mult mai greu decât să concurezi la *Jeopardy* și mult, mult mai greu decât să recunoști celulele canceroase într-o imagine.

O mașină de diagnosticare multifuncțională ar putea să pară doar un simplu pas logic de la algoritmii care se bazau pe imagini pentru depistarea cancerului pe care i-am întâlnit mai devreme, însă aceia aveau un mare avantaj: examinau celule reale care ar fi putut să cauzeze o anumită problemă. O mașină de diagnosticare, în schimb, primește informații în condițiile în care din fondul problemei sunt eliminate câteva etape. Poate că un pacient simte furnicături provocate de un spasm muscular produs de prinderea unui nerv din cauza ridicării unei greutăți prea mari. Sau poate că o persoană are sânge în scaun din cauza hemoroizilor apăruți în urma unei constipații provocate de alimentația necorespunzătoare. Un algoritm (sau un medic) trebuie să ia un singur simptom și să refacă traseul până când ajunge la diagnosticul corect. Asta trebuia să facă și Watson. Este o sarcină incredibil de dificilă.

Și mai există și alte probleme, pe deasupra!

Vă mai aduceți aminte de rețeaua neurală care deosebea un câine de un lup? Antrenarea sa fusese ușoară. Tot ce avuseseră de făcut programatorii fusese să găsească un teanc de fotografii etichetate „lup" sau „câine" și să le introducă în calculator. Setul de date era simplu și lipsit de ambiguități. „Numai că –

așa cum a declarat pentru *MIT Technology Review* Thomas Fuchs, specialist în patologie informatică*– într-un domeniu specializat al medicinei ai putea avea nevoie de specialiști instruiți timp de decenii care să nu facă nimic altceva decât să eticheteze corect informațiile introduse în computer."[40]

Această problemă ar putea fi depășită cu ajutorul unei întrebări foarte specializate (de exemplu, sortarea lamelelor cu țesut mamar care ajung la patologie în „complet benigne" și „foarte maligne"). Însă o mașină de diagnosticare universală, cum este Watson, ar avea nevoie să înțeleagă practic orice afecțiune posibilă. Ar fi nevoie de o armată de operatori umani incredibil de bine calificați, pregătiți să introducă în calculator informații despre diferiți pacienți și caracteristicile lor specifice pe o perioadă foarte lungă de timp. Și, dacă tot vorbim la modul general, patologii aceștia tind să aibă și alte lucruri de făcut – să salveze vieți, de exemplu.

Ajungem, astfel, la cea din urmă problemă. Cea care este cel mai dificil de depășit.

PROBLEMA DATELOR

Tamara Mills era doar un bebeluș când părinții ei au observat pentru prima oară că era ceva în neregulă cu felul în care respira. Când împlinise nouă luni, doctorii o diagnosticaseră ca fiind bolnavă de astm – o boală care afectează 5,4 milioane de oameni în Marea Britanie și 25 de milioane în SUA.[41] Cu toate că Tamara era mai tânără decât cei mai mulți suferinzi de astm, simptomele sale au fost perfect controlabile în primii ani de viață și fetița a crescut cam ca orice alt pușți cu această boală, petrecându-și copilăria jucându-se pe malul mării, în nordul Angliei (dar având tot timpul un inhalator la îndemână).

La opt ani, Tamara s-a îmbolnăvit de o formă urâtă de gripă porcină. De atunci, o infecție respiratorie s-a succedat după

* *Computational pathologist* (engl.) (n. tr.).

alta. Câteodată, în timpul atacurilor de astm, i se învineţeau buzele. Dar indiferent de câte ori Tamara şi mama ei se duceau la doctorii de la spitalul din apropiere, indiferent cât de des se plângeau părinţii ei că îşi termina rezerva de inhalatoare mai repede decât i se puteau prescrie[42], niciunul dintre doctori nu a trimis-o la un specialist.

Familia şi profesorii Tamarei, pe de altă parte, şi-au dat seama că lucrurile se agravau. După două atacuri aproape fatale care au dus la internarea fetei în spital, aceasta a fost scutită de orele de educaţie fizică de la şcoală. Când scările interioare au început să fie prea greu de urcat pentru ea, Tamara s-a dus să locuiască la bunici, într-un bungalow.

Pe 10 aprilie 2014, Tamara căzut victimă unei noi infecţii respiratorii. În noaptea aceea, bunicul ei a găsit-o chinuindu-se să respire. A chemat o ambulanţă şi a făcut tot ce i-a stat în putinţă să o ajute, cu două inhalatoare şi un tub de oxigen. Starea ei a continuat să se deterioreze. Tamara a murit câteva ore mai târziu, la numai 13 ani.

Astmul nu este, de obicei, o afecţiune fatală, dar, cu toate acestea, 1 200 de persoane de pe teritoriul Marii Britanii, dintre care 26 sunt copii, mor din cauza lui în fiecare an.[43] Se estimează că două treimi dintre aceste decese ar putea fi prevenite – aşa cum a fost şi cel al Tamarei. Numai că prevenirea depinde în întregime de identificarea semnelor de avertizare, în baza cărora se şi intervine.

Pe durata celor patru ani care au culminat cu cel din urmă atac, care i-a fost fatal, Tamara a fost la consultaţii la medicii de la spitalul din localitate de nu mai puţin de 47 de ori. Era evident că tratamentul ei nu dădea rezultate, dar, de fiecare dată când se ducea la un specialist în sănătate, acesta nu trata decât problema cea mai urgentă. Niciunul nu a luat în considerare imaginea de ansamblu. Niciunul nu a remarcat tiparul care reieşea la vizitele medicale, niciunul nu a observat

că starea fetei se deteriora constant, niciunul nu a sugerat că era momentul să încerce ceva nou.[44]

Pentru toate acestea exista un motiv. Indiferent dacă vă vine sau nu să credeți (și oricine locuiește în Anglia probabil că va crede), pentru NHS*, serviciul de sănătate național din Marea Britanie, corelarea evidențelor serviciilor medicale prestate nu este o procedură standard. Dacă te afli într-un spital din sistemul național de sănătate, medicii de acolo nu vor ști nimic despre vizitele pe care le-ai făcut la medicul de familie. Multe evidențe sunt încă ținute pe hârtie, ceea ce înseamnă că modul în care trec de la un medic la altul nu s-a schimbat de zeci de ani. Este unul dintre motivele pentru care sistemul de sănătate britanic deține titlul nu prea lăudabil de cel mai mare cumpărător de faxuri din lume.[45]

Oricât de nebunesc ar suna asta, noi, britanicii, nu suntem singurii. SUA au o puzderie de medici particulari și mari rețele de spitale care nu sunt conectate între ele; pe când alte țări, precum Germania, au început să introducă dosarele electronice de sănătate ale pacienților, deși acestea nu reprezintă, nici pe departe, ceva obișnuit peste tot în lume. Pentru Tamara Mills, lipsa unui istoric medical unic conectat la întregul sistem a însemnat că era imposibil pentru orice medic individual să înțeleagă pe deplin severitatea stării ei. Orice soluționare a acestei deficiențe va veni, din păcate, prea târziu pentru Tamara, dar rămâne o provocare uriașă pentru viitorul sistemului de sănătate. O mașină cum este Watson ar putea ajuta la salvarea unei mulțimi de Tamare, însă va reuși să găsească anumite tipare în informațiile pe care le deține numai dacă acestea sunt colectate, centralizate și conectate.

* National Health Service (engl.), Serviciul Medical Național, sistemul public de asistență medicală din Regatul Unit al Marii Britanii și Irlandei de Nord, care nu are o reputație bună în ceea ce privește protejarea datelor personale ale pacienților și angajaților săi (n. red.).

Există un contrast izbitor între seturile de date complete și detaliate deținute de brokerii de informații și seturile de date restrânse și necorelate din sistemul de sănătate. Pentru moment, informațiile medicale sunt vraiște. Chiar și atunci când istoricul nostru medical detaliat este stocat într-un singur loc (ceea ce adesea nu este cazul), informațiile însele pot lua atâtea forme, încât este practic imposibil ca acestea să fie corelate într-un mod utilizabil de către un algoritm. Sunt de luat în considerare radiografii, trebuie incluse rapoarte și fișe, rețete, observații; iar lista poate continua. Apoi mai este problema modului în care sunt înregistrate informațiile scrise. Trebuie să înțelegi acronimele și abrevierile, să descifrezi scrisul de mână, să identifici posibilitatea apariției unor erori umane. Toate acestea, înainte de a ajunge măcar la simptome. Oare această persoană s-a referit la faptul că are temperatură când a spus că suferă de o „răceală"? Sau „răceală" înseamnă că tușește? Pe această persoană chiar „o omoară" durerile de stomac? Sau simte doar puțină durere? Ideea este că medicina este foarte, foarte complicată și fiecare nivel de complexitate face ca informațiile să fie și mai greu de priceput pentru un computer.[46]

Cei de la IBM nu sunt singurii giganți economici *blue chip** care au avut de furcă din cauza problemei reprezentate de bazele de date dezordonate și nestructurate ale sistemului de sănătate. În 2016, DeepMind, divizia dedicată dezvoltării IA a companiei Google, a semnat un contract cu Royal Free NHS Trust din Londra. DeepMind a primit acces la informațiile medicale deținute de trei spitale municipale, în schimbul unei aplicații care i-ar fi ajutat pe medici să identifice insuficiențele renale acute. Inițial, se intenționa utilizarea algoritmilor inteligenți de învățare care să vină în sprijinul serviciilor

* Expresia *blue chip*, derivată din jargonul jucătorilor de pocher (care folosesc jetoane de diferite culori, dintre care cele albastre sunt cele mai valoroase), este o etichetă folosită pentru companiile publice considerate cele mai solide, mai eficiente și mai stabile de pe piață (n. tr.).

medicale, însă cercetătorii au descoperit că trebuiau să-şi limiteze ambiţiile şi să opteze pentru ceva mult mai simplu, deoarece informaţiile nu erau suficient de bune pentru îndeplinirea ţelurilor iniţiale.

Dincolo de provocările de natură strict practică, această colaborare dintre DeepMind şi NHS a dat naştere unei probleme încă şi mai controversate. Cercetătorii nu promiseseră decât să îi avertizeze pe medici despre prezenţa afecţiunilor renale, numai că Royal Free nu dispunea de un set de date despre rinichi pe care să li-l ofere. În schimb, DeepMind a primit acces deplin la *toate* informaţiile înregistrate în sistem, adică la antecedentele medicale ale unui număr de 1,6 milioane de pacienţi pe o perioadă de cinci ani.

Teoretic, o cantitate aşa de mare de informaţii ar putea ajuta la salvarea a nenumărate vieţi. Insuficienţele renale acute ucid o mie de persoane lunar, iar deţinerea unor informaţii care se întindeau pe o perioadă atât de îndelungată putea să ajute DeepMind să identifice tendinţe importante referitoare la evoluţia acestor afecţiuni. Mai mult decât atât, din moment ce bolile de rinichi se întâlnesc mai frecvent la persoane care suferă şi de alte maladii, un set de date cuprinzător ar uşura mult depistarea simptomelor şi a diverselor relaţii dintre variatele afecţiuni din prezent ale pacienţilor şi starea lor de sănătate viitoare.

În loc să fie întâmpinată cu entuziasm, anunţarea proiectului a stârnit însă un adevărat scandal. Şi nu fără temei. A-i oferi lui DeepMind accesul la *toate* informaţiile înregistrate de spitale însemna exact asta: *că avea acces la tot*. Compania Alphabet Inc. (care deţine şi Google) a aflat cine a fost internat în spital şi la ce dată. Cine a venit în vizită la pacienţi în perioada internării acestora. A aflat rezultatele rapoartelor patologice şi ale radiografiilor; cine a făcut întreruperi de sarcină, cine a suferit de depresie, până şi cine a fost diagnosticat cu HIV. Şi ştiţi ce este cel mai grav? Pacienţilor

însişi nu li s-a cerut niciodată consimţământul, nu li s-a oferit posibilitatea de a fi excluşi din studiu şi nici măcar nu li s-a spus că urmau să facă parte din acesta.[47]

Merită amintit că i s-a interzis companiei Google să utilizeze aceste informaţii în orice alt context al activităţii sale comerciale. Şi, cinstit vorbind, Google deţine în privinţa securităţii informaţiilor un palmares superior celui al NHS, ale cărui spitale au fost aduse în impas în 2017 de un virus informatic de tip *ransomware* lansat de Coreea de Nord, pentru că NHS încă mai folosea Windows XP.[48] Chiar şi aşa, este un motiv serios de îngrijorare faptul că o companie deja extrem de puternică, lider mondial în domeniul tehnologiei, a ajuns să aibă acces şi la informaţiile medicale ale unei persoane.

PROBLEME DE CONFIDENŢIALITATE

Haideţi să fim sinceri: Google nu duce lipsă de informaţii personale, intime chiar, despre fiecare dintre noi. Dar parcă simţim instinctiv ceva diferit – extrem de confidenţial – atunci când vine vorba despre informaţiile noastre medicale.

Pentru oricine se bucură de o stare de sănătate bună s-ar putea să nu fie atât de evident de ce: în definitiv, dacă ai avea de ales între a-ţi face publice informaţiile medicale sau istoricul navigării pe internet, ce ai prefera? Eu ştiu că l-aş alege, fără să stau pe gânduri, pe primul şi probabil că mulţi ar proceda la fel. Nu pentru că aş avea ceva deosebit de interesant de ascuns. Dar una este doar un instantaneu banal al fiinţei mele biologice, pe când cealaltă este o fereastră deschisă spre personalitatea mea.

Chiar dacă informaţiile despre starea de sănătate ar putea fi mai puţin stânjenitoare, Timandra Harkness, autoarea cărţii *Big Data: Does Size Matter?* şi prezentatoare a programului *Future Proofing* de la BBC Radio 4, susţine că acestea continuă să reprezinte un caz aparte:

Înainte de orice, informațiile despre starea de sănătate a multor persoane conțin o narațiune a vieții lor. Una din trei femei din Marea Britanie, de exemplu, a făcut o întrerupere de sarcină și ar putea exista persoane din viața lor care să nu fie la curent cu acest lucru.

Timandra subliniază și faptul că istoricul medical nu este important numai pentru tine:

Dacă o persoană deține informațiile tale genetice, atunci știe câte ceva și despre părinții, frații și copiii tăi. Iar când aceste informații sunt făcute publice, nu ai cum să mai scapi de asta. Nu-ți poți schimba structura biologică și nici nu o poți nega. Dacă examinează cineva un eșantion de ADN care îți aparține, nu poți schimba acest lucru. Poți face operații estetice la față, poți purta mănuși pentru a-ți ascunde amprentele, însă ADN-ul rămâne întotdeauna acolo. El este asociat cu persoana ta pentru totdeauna.

De ce este important acest lucru? Timandra mi-a relatat despre un *focus grup** pe care l-a condus în anul 2013, în care oameni obișnuiți erau întrebați ce îi preocupa cel mai mult în privința informațiilor lor medicale:

În ansamblu, oamenii nu erau îngrijorați de faptul că informațiile lor ar putea fi piratate sau furate. Îi preocupa faptul că s-ar putea face supoziții despre ei ca grup, care să fie apoi proiectate asupra lor ca indivizi.

Mai presus de orice, erau preocupați de modul în care aceste informații ar fi putea fi folosite împotriva lor:

Să presupunem că cineva și-a asociat cardul de fidelitate emis de un supermarket la istoricul medical. Această persoană s-ar putea duce să facă o operație de șold, iar medicul i-ar spune, de exemplu: „O, îmi pare rău, am constatat că ați

* Metodă de cercetare bazată pe discuții colective libere, care explorează o anumită întrebare sau un ansamblu de întrebări în cadrul unui grup bine definit (demografic, sociologic etc.) (n. tr.).

cumpărat o mare cantitate de pizza sau că ați cumpărat o mulțime de țigări, motiv pentru care mă tem că vom fi nevoiți să vă trimitem pe ultimul loc al listei de așteptare."

Este o temere foarte îndreptățită în Marea Britanie, unde anumite spitale publice ale NHS, secătuite cum sunt de fonduri, acordă deja prioritate nefumătorilor când vine vorba de realizarea unor operații de șold și genunchi.[49] Și există multe țări din lume în care pacienților obezi le pot fi respinse cererile de asigurare sau de tratament medical.[50]

Avem, aici, un fel de dilemă morală. Oamenii, ca specie, ar avea enorm de câștigat dacă le-ar permite algoritmilor să aibă acces direct la dosarele medicale algoritmi. Watson nu trebuie neapărat să rămână o simplă fantezie. Însă, pentru a-l transforma în realitate, va trebui să ne oferim informațiile medicale unor companii îndeajuns de bogate, încât să ne poarte prin corvoada încercărilor care se interpun între noi și acel doctor electronic fermecat. Iar, ori de câte ori renunțăm de bunăvoie la confidențialitate, ne jucăm periculos cu focul, fiind în pericol ca datele personale să ne fie eventual compromise, furate sau folosite împotriva noastră. Dumneavoastră sunteți pregătit să vă asumați acest risc? Aveți suficientă încredere în algoritmii aceștia și în beneficiile lor ca să-vă sacrificați intimitatea?

Sau, dacă se va ajunge aici, o să vă pese foarte tare de ea?

CÂND MOȘTENIREA GENETICĂ TE DĂ DE GOL

Francis Galton a fost un statistician și genetician din epoca victoriană, unul dintre cei mai remarcabili oameni ai generației sale și văr pe jumătate cu Charles Darwin. Multe dintre ideile sale au avut un efect profund asupra științei moderne și, nu în ultimul rând, opera sa a pus bazele statisticii moderne. Pentru aceasta îi datorăm lui Galton toată recunoștința noastră. (Din

păcate, el a fost destul de activ şi în mişcarea eugenică, aflată pe atunci în plin avânt, iar pentru asta, în mod cert, nu trebuie să îi fim recunoscători.)

Galton a vrut să studieze caracteristicile umane prin intermediul informaţiilor; chiar şi atunci, îşi dădea seama că era nevoie de o mulţime de date ca să afli ceva cu adevărat demn de interes. Dar mai ştia că oamenii nutreau o curiozitate insaţiabilă cu privire la propriul corp. Şi-a dat seama că – atunci când era stimulat în maniera potrivită – apetitul de nepotolit al oamenilor faţă de evaluarea lor de către un expert putea să depăşească nevoia de confidenţialitate. Mai mult decât atât, adesea erau dispuşi să plătească pentru plăcerea de a-şi satisface acest apetit.

Aşa se face că în 1884, când s-a organizat la Londra, sub patronajul reginei Victoria, o imensă expoziţie pentru a sărbători progresele Marii Britanii în domeniul sănătăţii, Galton şi-a văzut visul cu ochii. Şi-a instalat un stand la expoziţie pe cheltuiala sa – stand pe care l-a numit „Laborator antropometric" – în speranţa că va găsi câteva persoane printre milioanele de vizitatori dornice să plătească nişte bani pentru a fi măsurate.

A găsit chiar mai multe. Clienţii au format o coadă şi s-au îmbulzit la uşă, dornici să plătească un fleac de trei penny de persoană pentru a intra în laborator. Odată intraţi, puteau să-şi compare aptitudinile prin intermediul unor instrumente special concepute în acest scop, verificându-şi, printre altele, acuitatea vederii, forţa cu care puteau să tragă sau să strângă mâna şi viteza cu care trăgeau un pumn.[51] Laboratorul a fost atât de apreciat, încât Galton a trebuit să evalueze câte două persoane odată. (Şi-a dat repede seama însă că era mai bine să separe părinţii de copii în timpul testării pentru a evita pierderea de timp din cauza orgoliilor rănite. Într-un articol de ziar pe care l-a scris după eveniment, Galton comenta: „Celor vârstnici nu le plăcea să fie depăşiţi de cei tineri şi insistau să repete testele.")[52]

Indiferent cât de bine sau cât de rău se descurca fiecare persoană, rezultatele pe care le obținuse erau mâzgălite pe un cartonaș alb pe care îl putea păstra ca amintire. Adevăratul câștigător a fost, însă, Galton. Acesta a părăsit expoziția cu o copie completă a tuturor informațiilor – un set valoros de măsurători biometrice ale unui număr de 9 337 de persoane – și cu un profit frumușel pe deasupra.

Dacă trecem repede peste următorii 130 de ani, veți putea observa unele asemănări cu moda actuală a testărilor genetice. La un preț „de chilipir", de numai 149 de lire sterline, puteți trimite o probă de salivă companiei de genomică și biotehnologie 23andMe, pentru a primi în schimb un raport al trăsăturilor genetice, incluzând răspunsuri la întrebări de genul: ce tip de ceară de urechi produci? *Tu* ai cumva genele care duc la apariția unei monosprâncene?[53] Sau genele care te fac să strănuți când te uiți la soare?[54] Precum și la unele întrebări mai serioase: sunteți predispusă la cancer mamar? Aveți o predispoziție genetică pentru maladia Alzheimer?[55]

Între timp, compania a strâns, în acest mod ingenios, o uriașă bază de date cu informații genetice, care numără acum milioane de probe. Este exact ce au făcut și giganții internetului, numai că lor, în loc să le dăm ADN-ul nostru ca parte a târgului, le oferim informațiile cele mai personale pe care le deținem. Rezultă de aici o bază de date de pe urma căreia avem cu toții de câștigat. Aceasta constituie un bun de o valoare inestimabilă, căci ne va ajuta să înțelegem mai bine genomul uman. Oameni de știință, companii farmaceutice și organizații non-profit din întreaga lume așteaptă la rând să încheie parteneriate cu 23andMe pentru a depista anumite tipare în informațiile pe care aceasta le deține – atât cu, dar și fără ajutorul algoritmilor – în speranța găsirii unui răspuns la marile întrebări care ne afectează pe toți: care sunt cauzele ereditare ale diverselor afecțiuni? S-ar putea inventa medicamente noi, care să trateze persoanele care suferă de o

anumită boală? Există o cale mai bună de tratare a maladiei Parkinson?

Setul de date este valoros și într-un sens mai concret. Cu toate că fiecare cercetare întreprinsă aduce un beneficiu imens pentru societate, 23andMe nu face asta din bunătate și altruism. Dacă vă dați consimțământul (și 80% dintre clienți o fac), compania va vinde mai departe, pentru un profit frumușel, o versiune anonimizată a informațiilor dumneavoastră genetice partenerilor de cercetare mai sus-menționați.[56] Banii astfel câștigați nu sunt vreun bonus binevenit pentru companie, ci fac parte efectiv din planul ei de afaceri. Un membru din consiliul de administrație al 23andMe a declarat pentru *Fast Company*: „Jocul pe termen lung în acest caz nu este să facem bani din vânzarea de kituri, deși kiturile sunt indispensabile pentru obținerea informațiilor de bază." Iată ceva demn de reținut atunci când comandați un raport genetic comercial: nu dumneavoastră folosiți produsul. Dumneavoastră *sunteți* produsul.[57]

V-aș face un avertisment: eu, una, nu m-aș încrede prea mult în promisiunile lor de păstrare a anonimității. În 2005, un tânăr[58] conceput cu ajutorul unui donator de spermă rămas anonim a reușit să-l găsească și să-l identifice pe tatăl său biologic prin trimiterea unei probe de salivă la analiză și prin urmărirea indiciilor din propriul cod ADN.[59] Iar în 2013, un grup de oameni de știință a demonstrat, într-o lucrare foarte cunoscută, că milioane de oameni ar putea fi identificați după propriile gene, folosindu-se doar un calculator personal și câteva căutări inteligente pe internet.[60]

Și mai există un motiv pentru care s-ar putea să nu vă doriți ca ADN-ul dumneavoastră să fie înscris în vreo bază de date. Cu toate că există legi în vigoare care să protejeze oamenii împotriva celor mai urâte feluri de discriminare genetică – așa încât să nu ne îndreptăm chiar spre un viitor în care un Beethoven sau un Stephen Hawking să fie judecați

mai degrabă după predispozițiile lor genetice decât după talent –, aceste legi nu se aplică și în cazul asigurărilor de viață. Nimeni nu te poate sili să faci un test ADN dacă nu dorești acest lucru, însă în SUA asiguratorii au dreptul să te întrebe dacă ai făcut deja vreun test care îți calculează riscul de a dezvolta o anumită afecțiune, cum ar fi maladiile Parkinson, Alzheimer sau cancerul mamar, și pot să refuze încheierea unei asigurări de viață dacă răspunsul nu-i satisface. Iar în Marea Britanie asiguratorilor li se permite să țină cont de rezultatele testelor genetice pentru depistarea bolii Huntington (dacă suma acoperită depășește 500 000 de lire sterline).[61] Bineînțeles că ați putea încerca să mințiți și să susțineți că nu ați făcut niciodată un astfel de test, dar asta ar duce la invalidarea poliței de asigurare. Singura cale de a evita acest gen de discriminare este să nu faceți niciodată testul. Uneori, ignoranța chiar poate fi o binecuvântare.

Adevărul este că niciun set de date nu ar fi mai valoros pentru înțelegerea stării noastre de sănătate decât genomurile secvențiate a milioane de persoane. Cu toate acestea, nu cred că voi face un test genetic în viitorul apropiat. Însă (din fericire pentru societatea noastră, în ansamblu) milioane de oameni își *oferă* informațiile de bunăvoie. La ultima numărătoare, 23andMe avea peste două milioane de clienți genotipizați[62], iar MyHeritage, Ancestry.com – chiar și proiectul *Genografic** al National Geographic – aveau și mai multe milioane. Așadar, poate că aici nu există niciun fel de dilemă morală. În definitiv, piața și-a spus cuvântul: s-ar putea să nu merite ca în schimbul confidențialității informațiilor tale să ți se ofere șansa de a contribui la binele public, dar să descoperi că ești 25% viking merită cu prisosință.**

* Un proiect de cercetare a evoluției și migrației speciei umane lansat la 13 aprilie 2005 de Societatea National Geographic (n. tr.).

** De fapt, nimeni nu își poate da seama dacă cineva este sau nu viking, așa cum m-a informat pe larg bunul meu prieten, geneticianul Adam Rutherford. Am introdus această informație mai mult ca să-l

Bine, bine, recunosc: am glumit! Nimeni nu se așteaptă să se gândească, totuși, cineva în primul rând la marile provocări ale viitorului sistemului nostru de sănătate atunci când încearcă să ia decizia dacă să expedieze sau nu o probă de salivă pentru un test genetic. Este de înțeles dacă oamenii nu fac acest lucru – avem priorități diferite pentru noi înșine ca indivizi și pentru umanitate în ansamblu.

Dar asta ne trimite la un ultim aspect important. Dacă o mașină de diagnosticare capabilă să recomande tratamente medicale *poate* fi totuși construită, cui ar trebui ea să-i servească? Individului sau umanității? Pentru că vor fi momente în care ar putea fi nevoită să aleagă.

Imaginați-vă, de exemplu, că vă duceți la doctor din cauză că vă supără teribil tusea. Probabil că vă veți însănătoși de la sine, dar, dacă v-ar consulta o mașină – ca pacient –, ea ar putea dori să vă trimită să vă faceți o radiografie și un test de sânge, doar așa, pentru siguranță. Și probabil că v-ar prescrie antibiotice, dacă i le-ați solicita. Chiar dacă asta v-ar scurta suferința cu doar câteva zile, dacă unicul obiectiv ar fi propria sănătate și starea de bine, algoritmul ar putea considera că a meritat să vă prescrie tot acest tratament.

Dacă, însă, s-ar construi o mașină care să servească întregii populații, aceasta ar fi mult mai preocupată de problemele legate de rezistența la antibiotice. Atât timp cât nu v-ați afla într-un pericol iminent, disconfortul dumneavoastră de moment ar păli în importanță, iar algoritmul v-ar prescrie antibiotice numai dacă acest lucru ar fi absolut necesar. Un astfel de algoritm ar fi preocupat și de irosirea resurselor sau ar fi conștient de lungile liste de așteptare, astfel că nu v-ar

incit. Pentru a înțelege adevărata știință din spatele afirmației, citiți cartea sa, *A Brief History of Everyone Who Ever Lived: The Stories in Our Genes* (Weidenfeld & Nicolson, Londra, 2016) (n. a.).

trimite să faceți investigații suplimentare decât dacă ați avea și alte simptome care să denote ceva mai grav. Sincer vorbind, probabil că v-ar spune să luați o aspirină și să nu vă mai plângeți atât de milă.

În mod similar, o mașină care ar lucra în beneficiul tuturor ar putea să facă din „salvarea a cât mai multor vieți cu putință" o prioritate atunci când ar trebui să decidă cine să primească un transplant de organe. Ceea ce ar putea produce un plan de tratament diferit de cel al unei mașini care ar ține cont doar de interesul dumneavoastră personal.

O mașină care ar lucra pentru NHS sau pentru o companie de asigurări ar putea încerca să micșoreze costurile, dac-ar fi posibil, pe când una concepută să servească interesele unei companii farmaceutice ar putea avea ca obiectiv să promoveze utilizarea unui anumit medicament în locul altuia.

Situația medicinei este, cu siguranță, mai lipsită de tensiuni decât exemplele din sistemul de justiție penală din capitolul anterior. Aici nu avem apărare și acuzare. Toți cei din sistemul de sănătate au același deziderat: să facă pacientul să se simtă mai bine. Dar chiar și aici, fiecare parte implicată în proces are un set de obiective puțin diferit.

În oricare aspect al vieții s-ar introduce un algoritm, va trebui să se creeze întotdeauna un fel de echilibru. Între confidențialitate și binele public. Între individ și populație. Între diverse dificultăți și priorități. Nu este ușor să găsești o cale prin încâlceala de motivații, chiar și atunci când la capătul drumului așteaptă recompensa sigură a unui sistem de sănătate mai bun pentru toată lumea.

Dar este cu atât mai greu atunci când motivații concurente sunt ascunse vederii; când beneficiile unui algoritm sunt exagerate, iar riscurile – disimulate; sau când trebuie să te întrebi ce ți se cere să crezi și cine obține un profit de pe urma încrederii tale într-un anumit lucru.

Capitolul 5

AUTOMOBILELE AUTONOME

Soarele abia se ivise la orizont în dimineaţa zilei de 13 martie 2004, dar barul Slash X din inima deşertului Mojave deja gemea de oameni.[1] Localul se afla la periferia Barstowului, un orăşel situat între Los Angeles şi Las Vegas, situat în apropierea locului unde a fost filmată Uma Thurman ieşind dintr-un sicriu în filmul *Kill Bill II*.[2] Este un bar îndrăgit de cowboy şi de cei cărora le plac vehiculele de teren, dar, în acea dimineaţă de primăvară, reuşise să atragă un cu totul alt gen de clienţi. Stadionul improvizat care fusese construit afară, în praf, era înţesat de ingineri nebuni, de spectatori entuziasmaţi şi de pasionaţi nesăbuiţi, care împărtăşeau acelaşi vis: să fie primii oameni de pe planetă care asistau la câştigarea unei curse de către un vehicul fără şofer.

Cursa fusese organizată de US Defence Advanced Research Projects Agency*, DARPA (care mai era poreclită şi „departamentul de ştiinţă nebună al Pentagonului").[3] Agenţia era interesată de ceva timp de vehiculele fără şofer, şi pe bună dreptate: bombele artizanale şi atacurile ţintite asupra vehiculelor militare constituiau o cauză majoră a deceselor pe câmpul de luptă. Cu ceva timp în urmă, în acel an, îşi anunţase intenţia de a face autonome o treime din vehiculele forţelor militare la sol ale SUA până în anul 2015.[4]

* Agenţia pentru Proiecte de Cercetare Avansată în domeniul Apărării a SUA (n. tr.).

Până în acel moment, progresul fusese lent și costisitor. DARPA cheltuise în jur de o jumătate de miliard de dolari în peste 20 de ani, finanțând munca de cercetare din universități și companii, în speranța realizării acestei ambiții.[5] Și, la un moment dat, mai-marilor DARPA le-a venit o idee ingenioasă: ce-ar fi să organizeze un concurs? Ar invita în mod deschis toate persoanele interesate din țară, care să-și conceapă propriile vehicule fără șofer și să le pună să concureze una împotriva celeilalte pe o pistă de lungă distanță, oferind învingătorului un premiu de un milion de dolari.[6] Era primul eveniment de acest fel din lume și un mod rapid și ieftin de a oferi DARPA un avantaj în îndeplinirea obiectivului său.

Traseul cursei a fost amenajat pe ceva mai mult de 228 de kilometri, iar DARPA nu l-a conceput deloc pentru a fi ușor. Participanții trebuiau să se confrunte cu pante abrupte, bolovani, denivelări, șanțuri, porțiuni de teren accidentat și cactuși ciudați. Automobilele fără șofer trebuiau să parcurgă porțiuni cu zgură care, pe alocuri, aveau doar vreun metru și ceva lățime. Cu două ore înainte de start, organizatorii au oferit fiecărei echipe un CD cu coordonate GPS.[7] Acestea reprezentau 2 000 de puncte de trecere răspândite de-a lungul traseului asemenea firmiturilor de pâine – doar atât cât să le ofere vehiculelor o idee generală asupra direcției de mers, dar nu suficiente pentru a le ajuta să treacă peste obstacolele care le așteptau.

Provocarea era dificilă, însă 106 echipe curajoase s-au înscris la concurs în acel prim an, iar 15 au trecut de etapele de calificare, iar vehiculele lor au fost considerate suficient de sigure, încât să li se îngăduie să ajungă pe pistă. Printre acestea se numărau automobile care semănau cu buggyurile destinate mersului pe nisip, vehicule care semănau cu niște camioane gigantice și altele care păreau, mai degrabă, tancuri. Circulau zvonuri precum că un concurent își ipotecase casa ca să-și construiască automobilul fără șofer, pe când un altul lipise

două plăci de surf pe capotă ca să o facă să iasă în evidență. La întrecere a participat chiar și o motocicletă cu autobalans.[8]

În dimineața cursei, un șir de mașinării improvizate și care păreau cam șubrede s-au adunat la Slash X, alături de câteva mii de spectatori. Fără șoferi în interior, vehiculele s-au apropiat, pe rând, de linia de start; arătau de parcă ar fi venit, care mai de care, ori din *Mad Max,* ori din *Curse trăsnite*. Dar aspectul nu avea nicio importanță. Tot ce trebuiau să facă era un tur de pistă, fără nicio intervenție umană, în mai puțin de zece ore.

Însă lucrurile nu au mers deloc ca unse! Un automobil s-a răsturnat cu roțile în sus în zona de start și a trebuit retras din concurs.[9] Motoreta abia trecuse de linia de pornire, când s-a răsturnat pe o parte și a fost declarată eliminată din cursă. Un automobil a lovit un zid de beton după doar 45 de metri, altul s-a încurcat într-un gard din sârmă ghimpată, iar o mașinărie a fost prinsă între două tufișuri de mărăcini și – considerându-le obiecte imobile – s-a blocat pe manevre înainte și înapoi, înainte și înapoi, până când cineva a trebuit, totuși, să intervină.[10] Alte vehicule s-au izbit de bolovani și au intrat cu toată viteza în șanțuri. Osiile s-au rupt, cauciucurile s-au sfâșiat, caroseriile au zburat prin aer.[11] Peisajul din jurul barului Slash X a început să semene cu un cimitir de roboți.

Vehiculul cu cele mai bune rezultate, înscris la concurs de Universitatea Carnegie Mellon, a reușit să parcurgă impresionanta distanță de 11 kilometri înainte de a interpreta greșit un deal – moment în care cauciucurile au început să se rotească și, în lipsa intervenției umane, au continuat să se rotească până au luat foc.[12] Până la ora 11.00 dimineața, cursa s-a încheiat. Unul dintre organizatorii din cadrul DARPA s-a suit într-un elicopter și a zburat spre linia de sosire, să-i informeze pe jurnaliștii care așteptau că niciuna dintre mașinării nu avea să ajungă până acolo.[13]

Cursa fusese grea, plină de praf, zgomotoasă și distructivă – și se încheiase fără niciun câștigător. Toți oamenii din echipele

acelea munciseră un an întreg să creeze ceva care rezistase, în cel mai bun caz, doar câteva minute. Competiția s-a dovedit, însă, departe de a fi un dezastru. Concurența a dus la o explozie de noi idei, iar până la următorul Grand Challenge din 2005, tehnologia devenise aproape de nerecunoscut.

A doua oară, toți participanții, cu excepția unuia singur, au depășit distanța de 11 kilometri atinsă în 2004. Nu mai puțin de cinci vehicule diferite au reușit să parcurgă întreaga distanță a cursei, de 212 kilometri, fără nicio intervenție umană.[14]

Acum, cu puțin peste zece ani mai târziu, se acceptă pe scară largă ideea că viitorul aparține mijloacelor de transport fără șofer. Pe la sfârșitul anilor 2017, Philip Hammond, ministrul britanic de finanțe, a anunțat intenția guvernului de a avea pe șoselele britanice, până în anul 2021, automobile fără șofer complet autonome – fără însoțitor de siguranță la bord. Daimler a promis lansarea autovehiculelor fără șofer până în anul 2020,[15] Ford până în 2021[16], iar alți fabricanți au făcut propriile prognoze, similare cu cele citate.

Discuțiile din presă au trecut de la răsucirea pe toate părțile a întrebării dacă mașinile fără șofer urmează să apară, la dezbaterea provocărilor cărora vom fi nevoiți să le facem față atunci când vor fi peste tot. „Automobilul tău fără șofer ar trebui să lovească un pieton ca să-ți salveze ție viața?", întreba titlul unui articol din *New York Times* în iunie 2016.[17] Iar în noiembrie 2017: „Ce se va întâmpla în cazul animalelor omorâte pe șosea sau al primirii unei amenzi de circulație, atunci când vehiculele noastre vor deține controlul?"[18] Între timp, în ianuarie 2018, *Financial Times* avertiza: „Camioanele se îndreaptă spre un viitor din care vor lipsi șoferii: sindicatele avertizează că locurile de muncă a milioane de camionagii vor fi puse în pericol."[19]

Așadar, ce s-a schimbat? Cum s-a transformat tehnologia acestor vehicule, în numai câțiva ani, dintr-o însăilare șubredă într-una revoluționară și încrezătoarea în viitorul propriu? Și

putem, oare, nutri aşteptări rezonabile ca acest progres rapid să continue?

CE ESTE ÎN JURUL MEU?

Visul nostru despre un vehicul absolut autonom datează din epoca ştiinţifico-fantastică a *jet pack*-urilor*, a rachetelor cosmice, a costumelor spaţiale din folie de aluminiu şi a armelor laser. În 1939, la Expoziţia Universală din New York, compania General Motors a dezvăluit viziunea sa despre viitor. Vizitatorii s-au legat de un scaun echipat cu sistem audio instalat pe o bandă transportoare care i-a purtat într-un tur de 16 minute printr-o lume imaginară.[20] În spatele geamului, aceştia au văzut o machetă a visului companiei General Motors: super-autostrăzi care se întindeau de-a lungul şi de-a latul ţării, şosele care legau între ele ferme şi oraşe, străzi şi intersecţii – iar pe acestea goneau bolizi automatizaţi controlaţi prin unde radio care puteau să circule în siguranţă la viteze de până la 160 km/h. „Vi se pare cumva ciudat?" îi întreba prezentatorul. „Fantastic? Incredibil? Nu uitaţi, aceasta este lumea anilor 1960!"[21]

De-a lungul anilor au existat numeroase tentative de a transforma visul în realitate. General Motors a încercat acest lucru cu Firebird II, în anii 1950.[22] Cercetătorii britanici s-au străduit, în anii 1960, să adapteze un Citroën DS19 astfel încât acesta să comunice cu drumul pe care rula (undeva, între Slough şi Reading, încă mai găsiţi o porţiune de vreo 14 kilometri pe care este instalat un cablu electric, rămas de la experimentele lor).[23] Au urmat vehiculele din seria Navlab a celor de la Universitatea Carnegie Mellon din 1980; proiectul de un miliard de dolari

* Aparat de zbor individual prevăzut cu dispozitive de control al zborului ce pot fi manevrate cu ambele mâini şi înzestrat cu diferite sisteme de siguranţă (n. tr.).

și Eureka Prometheus, al Uniunii Europene din anii 1990.[24] Cu fiecare nouă încercare, visul construirii unei mașinării fără șofer părea ispititor de aproape, dar...

În aparență, construirea unui automobil fără șofer pare a fi relativ ușoară. Cei mai mulți oameni reușesc să stăpânească aptitudinile necesare pentru a conduce. Mai mult decât atât, există doar doi posibili vectori: viteza și direcția. E o chestiune de câtă accelerație să aplici și cât de mult să răsucești volanul. Cât de greu poate fi?

Așa cum a demonstrat primul Grand Challenge organizat de DARPA, construirea unui vehicul autonom este, de fapt, mult mai dificilă decât pare. Lucrurile se complică repede atunci când încerci să faci un algoritm să controleze o bucată mare și grea de metal care merge cu 100 km/h.

Să luăm rețelele neurale, de exemplu, care sunt utilizate cu mare succes pentru depistarea tumorilor din țesuturile mamare; ai putea crede că sunt foarte potrivite să ajute și tehnologia unui vehicul fără șofer „să vadă" mediul înconjurător. Prin 2004, rețelele neurale (aflate încă într-o formă rudimentară față de versiunile avansate ale celui mai recent stadiu de dezvoltare din zilele noastre) zbârnâiau deja înăuntrul prototipurilor vehiculelor fără șofer[25], încercând să înțeleagă ceva din ceea ce transmiteau camerele de filmat instalate pe capotele automobilelor. Este cert că se poate obține un mare volum de informații de valoare de la o cameră video. O rețea neurală poată să înțeleagă culoarea, textura, chiar și caracteristicile fizice ale scenei din fața sa – lucruri precum linii, curbe, margini și unghiuri. Întrebarea este: ce faci cu toate aceste informații atunci când dispui de ele?

Desigur, i-ai putea cere automobilului: „Rulează numai pe ceva care seamănă cu asfaltul". Dar acest lucru nu va fi de mare folos în deșert, unde șoselele nu sunt decât niște drumuri pline de praf. I-ai putea spune: „Rulează pe cel mai neted lucru din imaginea pe care o vezi", numai că, din păcate, cel mai neted

lucru este aproape întotdeauna cerul sau o clădire cu fațada din sticlă. Ai putea să gândești descrierea formei drumului în termeni mai abstracți: „Caută ceva cu două margini relativ drepte. Liniile ar trebui să fie la o distanță mare în partea de jos a imaginii, iar distanța să se micșoreze treptat în partea de sus." Pare destul de logic. Numai că, din păcate, la fel arată și un copac într-o fotografie. În general, nu este înțelept să încurajezi un vehicul să intre într-un copac.

Problema este că aceste camere video nu-ți pot oferi nicio idee despre proporții sau distanțe. Este un lucru pe care regizorii de film îl folosesc în permanență în avantajul lor – amintiți-vă numai scena de la începutul filmului *Star Wars* în care Star Destroyer își face încet-încet apariția pe fundalul negru precum cerneala al spațiului cosmic, planând dramatic spre partea de sus a cadrului. Ai impresia că este o namilă imensă, uriașă, când, în realitate, a fost filmat folosind o machetă nu mai mare de un metru lungime. Este o șmecherie care dă rezultate și arată grozav pe marele ecran. Dar, în cazul unui vehicul fără șofer pentru care două linii înguste, paralele, ar putea fi ori o șosea care continuă drept în față, ori trunchiul unui copac din apropiere, aprecierea corectă a distanței devine o chestiune de viață și de moarte.

Chiar dacă folosești mai multe camere video și combini în mod ingenios imaginile pentru a forma o imagine 3D despre lumea din jurul tău, apare o altă posibilă problemă din cauză că te bazezi prea mult pe rețelele neurale, așa cum a descoperit prin anii 1990 Dean Pomerleau, un om de știință de la Universitatea Carnegie Mellon. Acesta lucra la realizarea mașinăriei numite ALVINN*, pe care o instruia cum să înțeleagă mediul înconjurător din acțiunile unui șofer uman. Pomerleau și alții se așezau la volan și îl scoteau pe ALVINN la plimbări lungi, înregistrând tot ce făceau în timp ce conduceau

* Autonomous Land Vehicle In a Neural Network (engl.), „Vehicul terestru autonom într-o rețea neurală" (n. tr.).

vehiculul. În felul acesta a fost creat setul de date de antrenament din care rețelele lor neurale trebuiau să învețe o lecție: „condu oriunde ar conduce și un om, evită orice altceva".[26]

La început, totul a funcționat de minune. După antrenament, ALVINN a reușit să navigheze cu ușurință, de unul singur, pe un drum simplu. Numai că ALVINN a întâlnit un pod și totul a luat-o razna. Dintr-odată, a intrat într-un viraj periculos, iar Pomerleau a trebuit să apuce în grabă volanul, să prevină un accident.

După săptămâni întregi în care n-a făcut decât să analizeze informațiile despre incident, și-a dat seama care fusese problema: toate drumurile pe care fusese antrenat ALVINN erau mărginite de iarbă. Asemenea rețelelor neurale din capitolul „Medicina" care clasau câinii husky în funcție de zăpada din fotografii, și rețeaua neurală a lui ALVINN folosise iarba ca indicator principal pentru stabilirea direcției în care să conducă. În momentul în care iarba a dispărut, automobilul nu a mai avut habar ce să facă.

Spre deosebire de camerele video, laserele chiar *pot* să măsoare distanțele. Vehiculele care folosesc sistemul numit LiDAR (Light Detection and Ranging*, utilizat pentru prima oară în timpul celui de-al doilea Grand Challenge organizat de DARPA în 2005) emit un foton dintr-un laser, cronometrează de cât timp are nevoie acesta pentru a se lovi de un obstacol și a se întoarce înapoi, obținând astfel o estimare adecvată a distanței la care se află obstacolul respectiv. Dar nu sunt numai vești bune: LiDAR nu ajută în niciun fel la depistarea texturii sau a culorii unui obiect, este inutil când se pune problema să interpreteze semnele de circulație și nu este foarte performant pe distanțe lungi. Radarul, pe de altă parte – care face același lucru, dar folosind unde radio – se descurcă

* Numit și *radar laser*, este un sistem similar radarului, care utilizează pentru emisie radiația laser, cu lungimi de undă în spectrul optic (inclusiv ultraviolet și infraroșu) (n. tr.).

de minune, indiferent de condițiile atmosferice, poate să detecteze obstacole aflate la mare distanță, chiar și să vadă prin anumite materiale, însă este absolut inutil în furnizarea vreunui detaliu despre forma sau structura obstacolului.

De sine stătătoare, niciunul dintre aceste aparate care devin surse de date – camera video, sistemul LiDAR sau radarul – nu este destul de performant pentru a înțelege ce se întâmplă în jurul unui vehicul. Pentru construirea cu succes a unui vehicul fără șofer, șmecheria este să le combini. Ceea ce ar fi o sarcină relativ ușoară dacă toate instrumentele s-ar pune de acord asupra a ceea ce văd, numai că este mult mai dificil atunci când nu fac acest lucru.

Amintiți-vă numai de tufișul de mărăcini care a pus în încurcătură unul dintre autovehicule la primul Grand Challenge organizat de DARPA și imaginați-vă că automobilul dumneavoastră fără șofer s-ar găsi într-o situație asemănătoare. LiDAR transmite că aveți un obstacol în față. Camera video este și ea de acord. Radarul, care poate să vadă prin tufișul firav de mărăcini, comunică însă că nu aveți de ce să vă îngrijorați. Pe care dintre senzori ar trebui să-l creadă algoritmul dumneavoastră?

Ce s-ar întâmpla în cazul în care ar prevala camera video? Închipuiți-vă că vă iese în cale un autocamion mare și alb, într-o zi noroasă. De data asta, LiDAR-ul și radarul cad de acord că trebuie frânat, numai că pe fundalul cerului albicios și mohorât camera video nu poate să vadă nimic care să reprezinte un pericol.

Iar dacă toate acestea nu vi se par destul de dificile, atunci aflați că mai există o problemă. Nu trebuie să vă faceți griji numai din cauză că senzorii ar putea să interpreteze greșit mediul înconjurător, trebuie să țineți cont și de faptul că s-ar putea să îi măsoare greșit.

Poate că ați observat cerculețul acela albastru din Google Maps care încercuiește locul în care vă aflați; el se află acolo ca să indice o posibilă eroare în redarea semnalelor GPS. Uneori,

cercul albastru va fi mic și va marca exact poziția în care sunteți; alteori, el va acoperi un spațiu mai mare și va fi centrat într-un loc greșit. De cele mai multe ori, asta nu contează prea mult. Știm unde suntem și putem să respingem informația inexactă. Dar un vehicul fără șofer nu are o confirmare de la fața locului a poziției în care se află. Atunci când rulează pe o singură bandă de autostradă, mai îngustă de patru metri, nu se poate baza numai pe GPS pentru a recunoaște exact unde se află.

GPS-ul nu este singurul care oferă coordonate ce pot fi incerte. Fiecare măsurătoare realizată de un automobil va avea o marjă de eroare oarecare: înregistrările radarului, rotația pneurilor, inerția vehiculului... Nimic nu este niciodată 100% demn de încredere. Mai mult decât atât, condițiile atmosferice înrăutățesc lucrurile: ploaia afectează LiDAR-ul[27], razele puternice ale soarelui pot să perturbe funcționarea camerelor video,[28] iar drumurile lungi și accidentate distrug accelerometrele.[29]

Așa că, în cele din urmă, te vei alege doar cu o harababură de semnale. Întrebări care păreau simple („Unde ești?" „Ce se află în jurul tău?" „Ce ar trebui să faci acum?") devin uimitor de dificile. Este aproape imposibil să știi ce să mai crezi.

Aproape imposibil. Nu chiar 100% imposibil.

Pentru că, din fericire, îți poți croi un drum prin tot haosul acesta – există o cale de a face supoziții rezonabile într-o lume dezordonată, haotică. Totul se reduce la o formulă matematică extraordinar de importantă, cunoscută drept „teorema lui Bayes".

MAREA BISERICĂ A REVERENDULUI BAYES

Nu este nicio exagerare să afirmăm că teorema lui Bayes este una dintre cele mai influente idei din istorie. Ea trezește un entuziasm aproape reverențios printre oameni de știință, specialiști în învățarea automată și statisticieni, deopotrivă. Ideea de bază este, însă, extraordinar de simplă. Atât de

incredibil de simplă, de fapt, încât, la o primă privire, ați putea crede că nu face decât să expună ceva ce este deja evident.

Dați-mi voie să ilustrez această idee cu un exemplu cât de poate de banal.

Imaginați-vă că stați pe un scaun și luați masa la un restaurant. La un moment dat, persoana care vă însoțește se apleacă și vă șoptește că a observat că Lady Gaga ia prânzul la masa vecină.

Înainte de a vă întoarce să priviți, vă veți întreba, desigur, câtă încredere să acordați spuselor prietenului dumneavoastră. Așa că veți lua în considerare toate cunoștințele anterioare: calitatea restaurantului, poate, distanța până la locuința artistei din Malibu sau chiar acuitatea vizuală a prietenului... Ceva de genul acesta. Dac-ar trebui să ofer o definiție: este o convingere care poate fi cuantificată. Ca un fel de probabilitate.

Și, în timp ce vă întoarceți să o priviți pe femeia în cauză, veți folosi, în mod automat, fiecare dovadă de care dispuneți ca să vă actualizați nivelul de încredere în ipoteza prietenului dumneavoastră. Poate că părul blond platinat corespunde cu ceea ce v-ați aștepta să vedeți la Lady Gaga, astfel încât încrederea dumneavoastră sporește. Însă faptul că stă la masă fără niciun bodyguard lângă ea nu mai corespunde, astfel încât încrederea scade. Ideea este că fiecare detaliu nou contribuie cu ceva la evaluarea generală pe care o întreprindeți.

Asta este tot ce face teorema lui Bayes: oferă un mod sistematic de a ne actualiza nivelul de încredere într-o ipoteză pe baza unor dovezi.[30] Ea acceptă faptul că nu poți fi niciodată absolut sigur de teoria la care te gândești, dar îți permite să faci supoziția cea mai bună pe baza informațiilor de care dispui. Așadar, când vă dați seama că femeia de la masa vecină poartă o rochie confecționată din carne* – o alegere stilistică pe care este puțin probabil să o întâlniți din întâmplare la vreo ființă care

* Lady Gaga este artista care, în 2010, a apărut la gala de decernare a premiilor Video Music Awards îmbrăcată într-o rochie confecționată din carne (n. tr.).

să nu fie Lady Gaga –, acest lucru ar putea fi suficient pentru ca încrederea dumneavoastră să depășească pragul și să trageți concluzia că în acel restaurant se află, într-adevăr, Lady Gaga.

Însă teorema lui Bayes nu este numai o ecuație care definește modul în care oamenii iau decizii. Este mult mai importantă de atât. Pentru a o cita pe Sharon Bertsch McGrayne, autoarea cărții *The Theory That Would Not Die*:

> [Teorema lui] Bayes vine în contradicție cu profunda noastră convingere că știința modernă are nevoie de obiectivitate și de precizie.[31]

Oferind un mecanism de măsurare a încrederii pe care o ai într-un anumit lucru, teorema lui Bayes îți permite să tragi concluzii de bun-simț din observații sumare, din informații dezordonate, incomplete și aproximative – ba chiar și din neștiința ta.

Teorema lui Bayes nu există numai ca să confirme intuițiile existente. Se dovedește că, dacă ești nevoit să-ți cuantifici convingerile într-un anumit lucru, se ajunge adesea la concluzii paradoxale, care par să contrazică intuiția. Teorema lui Bayes explică de ce mai mulți bărbați decât femei sunt identificați în mod greșit drept viitori ucigași (vezi exemplul din capitolul „Justiția"). Și tot teorema lui Bayes explică de ce – chiar dacă ai fost diagnosticată ca suferind de cancer mamar –, dat fiind nivelul semnificativ de eroare al testelor, înseamnă că probabil nu ai așa ceva (vezi capitolul „Medicina"). Teorema lui Bayes este, în toate domeniile științei, un instrument puternic pentru selectarea și înțelegerea a ceea ce cunoaștem de fapt.

Însă modul de gândire bayesian își arată cu adevărat valoarea atunci când încerci să iei simultan în considerare mai mult de o singură ipoteză – de exemplu, atunci când te străduiești să îți dai seama ce nu este în regulă cu un pacient

pe baza simptomelor pe care le manifestă* sau să găsești locul unde se află un vehicul fără șofer pe baza indicațiilor oferite de senzorii acestuia. Teoretic, orice boală sau orice punct de pe o hartă ar putea să reprezinte un adevăr întemeiat, bazat pe date observate. Tot ce ai de făcut este să evaluezi elementele ca să decizi care are șansele cele mai mari să fie corect.

Iar din acest moment, găsirea poziției unui automobil fără șofer se dovedește a fi destul de asemănătoare cu ceea ce l-a pus în dificultate pe Thomas Bayes, pastorul prezbiterian și talentatul matematician britanic al cărui nume îl poartă teorema. Pe la mijlocul anilor 1700, Bayes a scris un eseu în care a inclus detaliile unui joc pe care îl inventase pentru explicarea problemei. Era ceva de genul acesta:[32]

Imaginați-vă că stați cu spatele la o masă pătrată. Fără să vedeți, eu arunc pe masă o minge roșie.

Dumneavoastră trebuie să ghiciți unde a aterizat aceasta. Nu va fi ușor: fără nicio informație de la care să puteți pleca, nu există nicio modalitate reală de a ști unde ar putea fi mingea pe masă.

Atunci, ca să vă ajut să ghiciți, arunc o a doua minge, de culoare diferită, pe aceeași masă.

Trebuie să faceți același lucru, și anume să stabiliți unde se află prima minge, cea roșie, însă acum vă voi spune unde ajunge pe masă cea de-a doua minge față de prima: dacă este în fața ei, în spatele ei, la stânga sau la dreapta mingii roșii. Iar dumneavoastră vă veți putea corecta evaluarea inițială.

Apoi vom repeta: voi arunca a treia, a patra și a cincea minge pe masă și de fiecare dată vă voi spune unde aterizează

* Watson, calculatorul IBM despre care am discutat în capitolul „Medicina", folosește din plin așa-numita „inferență bayesiană". Vezi postarea „Watson and Healthcare", on-line la adresa: www.ibm.com/developerworks/library/os-ind-watson/ (accesată cel mai recent la 19 martie 2019) (n. a.).

fiecare dintre ele în raport cu prima minge, cea roşie – a cărei poziţie pe masă încercaţi să o ghiciţi.

Cu cât voi arunca mai multe mingi şi cu cât vă voi oferi mai multe informaţii, cu atât imaginea poziţiei mingii roşii ar trebui să vă fie tot mai clară.

Nu veţi fi niciodată absolut sigur unde se află cu exactitate, dar vă puteţi actualiza nivelul de încredere cu privire la poziţia ei până când veţi sfârşi prin a găsi un răspuns care să vă satisfacă pe deplin.

Într-un fel, poziţia reală a vehiculului fără şofer este asemănătoare cu aceea a mingii roşii.

În locul unei persoane care stă cu spatele la masă, avem de-a face cu un algoritm care încearcă să evalueze cu precizie unde se află vehiculul la un moment dat, iar în locul celorlalte mingii aruncate pe masă, avem sursele de informaţii: GPS-ul, măsurarea inerţiei şi aşa mai departe. Niciuna dintre acestea nu-i spune algoritmului unde se află automobilul cu exactitate, dar fiecare adaugă încă puţină informaţie, pe care algoritmul o poate folosi pentru a-şi actualiza nivelul de încredere.

Este o şmecherie cunoscută drept „inferenţă probabilistică" – folosirea informaţiilor (plus teorema lui Bayes) pentru a face o inferenţă despre poziţia reală a obiectului. Şi, „ambalată" în mod corespunzător, devine doar încă unul dintre algoritmii de învăţare automată.

Pe la începutul acestui mileniu, inginerii acumulaseră deja suficientă experienţă practică cu rachetele de croazieră, navetele spaţiale şi aparatele de zbor, ca să ştie cum să abordeze problema poziţiei. A face un vehicul fără şofer să răspundă la întrebarea „Unde mă aflu?" nu era încă o însărcinare banală, dar, cu puţină gândire bayesiană, măcar era un lucru realizabil.

De la adevăratul cimitir de roboţi de la Grand Challenge din 2004, la uimitorul triumf al tehnologiei de la evenimentul din 2005 – în timpul căruia cinci vehicule diferite au reuşit să

parcurgă peste 160 de kilometri fără nicio intervenție umană – multe dintre cele mai mari salturi înainte au avut loc datorită lui Bayes. Algoritmii care se bazau pe ideile bayesiene au ajutat la soluționarea celorlalte întrebări la care mai trebuia să răspundă un autovehicul fără șofer: „Ce se află în jurul meu?" și „Ce ar trebui să fac acum?"*

AȘADAR, AUTOMOBILUL TĂU FĂRĂ ȘOFER AR TREBUI SĂ LOVEASCĂ UN PIETON CA SĂ-ȚI SALVEZE VIAȚA?

Să ne oprim o clipă și să ne gândim la cea de-a doua întrebare. Pentru că în anul 2016, pe la începutul toamnei, retras într-un colț mai liniștit al sălii de expoziție, altfel plină de agitație, a Salonului Auto de la Paris, un purtător de cuvânt al companiei Mercedes-Benz a făcut un anunț destul de ieșit din comun pe această temă. Christoph von Hugo, directorul departamentului de sisteme de asistență la volan și sisteme de siguranță al companiei, a fost întrebat, într-un interviu, ce ar trebui să facă un vehicul fără șofer în cazul unui accident. Iar el a răspuns:

> Dacă știi că poți salva cel puțin o persoană, atunci salveaz-o măcar pe aceea.[33]

Este doar o chestiune de logică pură, ați putea crede. În niciun caz nu merită să devină subiectul unei știri de pe prima pagină a ziarelor.

Numai că Hugo nu fusese întrebat despre cine știe ce accident de demult. El fusese pus la încercare și rugat să ofere un răspuns la un bine-cunoscut experiment de gândire,

* Câștigătoarea cursei din 2005, care fusese creată de o echipă de la Universitatea Stanford, a fost descrisă destul de bine de matematicianul Pesri Diaconis, de la aceeași universitate, care a declarat: „Fiecare șurub al vehiculului aceluia era bayesian" (n. a.).

conceput în anii 1960, care descria un tip aparte de coliziune. Reporterul se referea la o dilemă morală ciudată, care te forțează să alegi între două rele, cunoscută drept „problema vagonului", după tramvaiul scăpat de sub control descris în formularea inițială. În cazul automobilelor autonome, ea sună astfel:

> Imaginați-vă că, peste câțiva ani, vă aflați într-un vehicul autonom și conduceți fără nicio grijă pe o stradă dintr-un oraș. În fața dumneavoastră, un semafor indică roșu, însă o defecțiune mecanică a automobilului vă pune în imposibilitatea de a-l opri. O coliziune este inevitabilă, însă vehiculul poate să facă o alegere: să vireze, părăsind drumul, și să intre într-un zid de beton, ceea ce va duce la moartea sigură a tuturor celor din vehicul. Sau să-și continue traiectoria, salvând viețile tuturor celor din interior, dar să omoare pietonii care traversează strada. Cum ar trebui programat vehiculul să acționeze? Cum decideți cine ar trebui să moară?

Fără îndoială că aveți propria opinie. Poate credeți că mașina ar trebui să încerce să salveze cât mai multe vieți cu putință. Sau poate considerați că porunca biblică „să nu ucizi" ar trebui să prevaleze asupra oricăror altor considerente, lăsându-l pe cel aflat în mașină să suporte consecințele.*

Hugo a fost lipsit de echivoc în privința poziției companiei Mercedes. „Trebuie salvat cel din mașină!" Și a continuat: „Dacă știi cu siguranță că o moarte poate fi prevenită, atunci aceea este prioritatea."**

* În presă, de la *New York Times* la *Mail on Sunday*, au apărut numeroase versiuni ale acestui scenariu. „Dacă pietonul este o bunicuță de 90 de ani? Dacă este un copilaș? Dacă în automobil s-ar afla un laureat al Premiului Nobel?" Toate au la bază aceeași dilemă (n. a.).

** Citatele din declarația lui Cristoph von Hugo au fost preluate din articolul postat on-line la adresa: http://www.automarket.ro/stiri/mercedes-a-rezolvat-problema-morala-masinile-autonome-vor-73753.html (accesată cel mai recent la 15 martie 2019) (n. tr.).

În zilele care au urmat interviului, internetul a fost inundat de articole care incriminau atitudinea companiei. „Mașinile lor vor acționa foarte asemănător cu modul în care ar face-o un răzgâiat șofer european al unei mașini de lux"[34], scria autorul unui articol. Într-adevăr, într-un sondaj publicat în revista *Science* în aceeași vară,[35] 76% dintre respondenți simțeau că ar fi un gest mai etic din partea vehiculelor autonome ca acestea să salveze cât mai multe vieți cu putință, omorând astfel persoanele din interiorul lor. Mercedes nimerise de partea opusă a opiniei generale.

Așa să fie, oare? Pentru că atunci când, în același sondaj, participanții au fost întrebați dacă ei înșiși chiar *ar cumpăra* o mașină care i-ar ucide dacă împrejurările ar cere-o, aceștia au devenit dintr-odată reticenți în a se sacrifica pe ei înșiși pentru binele general.

Aceasta este o dilemă morală care face ca opiniile să fie împărțite – și nu numai în privința răspunsului pe care oamenii îl consideră cel mai nimerit. Ca experiment de gândire, rămâne favoritul de necontestat al reporterilor specializați în tehnologie și al altor jurnaliști, însă toți specialiștii în automobile autonome pe care i-am intervievat au dat ochii peste cap imediat ce am amintit de problema vagonului. Eu, personal, încă mai am o slăbiciune față de ea. Simplitatea sa ne silește să recunoaștem ceva important despre vehiculele fără șofer, să reflectăm la ceea ce simțim față de un algoritm care face o judecată de valoare în privința vieții noastre și a altor persoane. La temelia acestei noi tehnologii – așa cum este cazul cu aproape toți algoritmii – se află aspecte care privesc puterea, așteptările, controlul și delegarea responsabilității. Precum și dacă ne putem aștepta ca tehnologia pe care o inventăm să ni se potrivească nouă și nu viceversa. Înțeleg însă și reacția rezervată pe care o primesc din partea celor care se ocupă de crearea automobilelor autonome. Ei știu, mai bine decât oricine altcineva, cât de

departe suntem de momentul în care chiar va trebui să ne facem griji din cauza problemei vagonului.

ÎNCĂLCAREA REGULILOR DE CIRCULAȚIE

Teorema lui Bayes și puterea probabilității au propulsat o mare parte a inovației în domeniul automobilelor autonome încă de la prima competiție organizată de DARPA. L-am întrebat pe Paul Newman, profesor de robotică la Universitatea Oxford și fondatorul Oxbotica, o companie care construiește vehicule fără șofer și le testează pe șoselele din Marea Britanie, cum se comportau cele mai recente automobile autonome ale sale, iar el mi-a explicat următoarele:

> Sunt multe, multe milioane de linii de cod, dar am reușit să încadrez toată chestia drept o inferență probabilistică. Pe toată![36]

Dar în timp ce inferența bayesiană se îndreaptă oarecum spre explicarea modului în care este posibilă existența mașinilor fără șofer, aceasta arată totodată și că autonomia deplină, lipsită de orice intervenție a unui șofer uman, este o problemă foarte, foarte dificil de rezolvat.

Închipuiți-vă ceea ce sugerează Paul Newman: „ai două vehicule care se apropie cu viteză unul de celălalt" – rulând, să spunem, în direcții opuse pe o șosea A* în curbă lină. Un șofer uman se va simți confortabil în acest scenariu, știind că mașina cealaltă va rămâne pe banda ei de mers și că va trece de el în siguranță, la vreo doi metri de marginea marcajului. „Însă, pentru câteva minute – explică Newman – totul arată ca și cum mașinile vor intra în coliziune." Cum înveți un automobil autonom să nu intre în panică într-o

* În Marea Britanie, o șosea A este o cale de circulație principală, mai îngustă decât o autostradă, dar mai lată și mai dreaptă decât o șosea B (n. tr.).

asemenea situație? Nu-ți dorești ca vehiculul să devieze de pe șosea, încercând să evite o coliziune care nu va avea loc niciodată. De asemenea, nu ți-ai dori nici să rămână nepăsător dacă într-adevăr te găsești în pragul unei coliziuni frontale. Aduceți-vă aminte că aceste mașini nu fac decât să aprecieze cu aproximație ce au de făcut. Cum să le facem să ghicească de fiecare dată corect? Asta, spune Newman, „este o problemă foarte, foarte dificilă."

Este o problemă care i-a pus mult timp în impas pe specialiști, însă ea are o soluție. Trucul este să incluzi din construcție un model care să știe cum se vor comporta ceilalți șoferi (cei care sunt sănătoși la minte). Din nefericire, nu se poate spune același lucru și despre alte scenarii de conducere particulare.

Newman explică: „Cea mai dificilă parte o constituie problemele care apar în timp ce șofezi și care nu au nimic de-a face cu conducerea autovehiculului." De exemplu, să înveți un algoritm să înțeleagă că trebuie să fie extrem de atent atunci când aude melodia unei mașini din care se vinde înghețată sau când trece pe lângă un grup de copii care se joacă liniștiți cu mingea pe trotuar. Sau să recunoască țopăitul derutant al unui cangur – o chestiune pe care, în momentul în care am scris această carte, compania Volvo a recunoscut că se chinuia să o soluționeze.[37] Probabil că nu este o problemă chiar atât de importantă în rusticul Surrey, dar este ceva ce mașinile trebuie să-și însușească pentru a corespunde cerințelor de pe drumurile publice din Australia.

Iată ceva încă și mai greu de realizat: cum înveți o mașină că, uneori, ar trebui să încalce regulile de circulație? Ce ar fi dacă staționezi la semafor și cineva apare brusc în fața mașinii și îți face frenetic semn să avansezi? Sau dacă o ambulanță cu semnalele luminoase aprinse încearcă să străbată o stradă îngustă, iar tu trebuie să te urci pe trotuar ca s-o lași să treacă? Sau dacă o cisternă care transportă petrol s-a răsturnat pe un drum de țară, iar tu trebuie să părăsești locul cu orice preț?

„Nicio situație de acest gen nu figurează în Codul rutier", subliniază Newman în mod corect. Și, totuși, o mașină cu adevărat autonomă trebuie să știe cum să facă față tuturor acestor împrejurări, dacă dorim ca aceasta să funcționeze fără *absolut* nicio intervenție umană. Chiar și în situații de urgență.

Ceea ce nu înseamnă că aceste probleme sunt insurmontabile. „Nu cred că există vreun nivel de inteligență la care să nu reușim să aducem o mașină", mi-a spus Newman. „Singura întrebare care se pune este când vom reuși acest lucru."

Din păcate, răspunsul la dilema de mai sus este: probabil că nu în viitorul apropiat. Visul realizării unui automobil autonom, care ne dorim cu toții să se îndeplinească, ar putea fi cu mult mai îndepărtat decât credem.

Asta, pentru că mai există un nivel de dificultate pe care trebuie să-l depășim atunci când încercăm să construim această fantasmagorie științifico-fantastică de vehicul fără șofer, fără volan, care te duce oriunde și face orice. Iar el depășește cu mult provocările tehnologice. Un automobil complet autonom trebuie să se descurce și cu problema spinoasă reprezentată de factorul uman.

Jack Stilgoe, sociolog la University College London și specialist în impactul social al tehnologiei, explică:

> Oamenii sunt răutăcioși. Ei sunt agenți activi, nu doar elemente pasive într-un scenariu.[38]

Imaginați-vă, pentru o clipă, o lume în care există vehicule cu adevărat și perfect autonome. Prima regulă a algoritmilor de la bord va fi să evite coliziunile ori de câte ori este posibil. Iar aceasta schimbă relațiile dintre participanții la trafic de pe drumuri. Dacă stai în fața unui automobil fără șofer, acesta trebuie să oprească. Dacă intri în fața unui vehicul autonom la o intersecție, acesta trebuie să-ți cedeze trecerea.

Cum a declarat unul dintre participanții la un *focus group* organizat în 2016 la London School of Economics: „Pur și

simplu, le vei lua fața. Se vor opri, iar tu vei trece pe lângă ele." În traducere: aceste vehicule vor putea fi intimidate.

Stilgoe este de aceeași părere:

> Persoanele care până acum au fost oarecum lipsite de putere pe șosele, cum ar fi bicicliștii, ar putea începe să pedaleze foarte încet în fața automobilelor autonome, știind că acestea nu vor putea face niciodată niciun gest agresiv.

Soluționarea problemei ar putea însemna adoptarea unor reguli mai stricte împotriva acelora care abuzează de condiția lor de ciclişti sau de pietoni. Este o măsură care a mai fost luată și altădată, bineînțeles: gândiți-vă numai la interzicerea traversării prin locuri nepermise. Sau ar putea însemna eliminarea forțată a orice altceva mai circulă pe șosele – așa cum s-a întâmplat în momentul apariției autoturismelor –, motiv pentru care, în prezent, nu mai întâlnim biciclete, cai, care, căruțe sau pietoni pe o autostradă.

Dacă ne dorim vehicule complet autonome, atunci, aproape cu certitudine, vom fi nevoiți să luăm măsuri asemănătoare și pentru a limita numărul șoferilor agresivi, al mașinilor de înghețată, al copiilor care se mai joacă încă în stradă, al semnelor rutiere care anunță „lucrări pe partea carosabilă", dar și numărul pietonilor dificili, al ambulanțelor, al bicicliștilor, al scuterelor și al oricui altcuiva care face problema autonomiei atât de dificilă. Toate ar fi bune și frumoase, numai că viziunea aceasta diferă puțin de modul în care ne sunt prezentate nouă lucrurile în prezent, pentru că, așa cum mi-a declarat Stilgoe:

> Toate cuvântările frumoase despre autonomie și mijloacele de transport au de-a face, mai degrabă, cu modul în care să *nu* schimbăm lumea. Cu alte cuvinte, cum să păstrăm lumea așa cum este, dar să construim și să permitem unui robot să fie tot atât de priceput ca o ființă umană, apoi încă și mai priceput decât aceasta la conducerea unei mașini. Iar eu cred că asta este o mare prostie.

„Dar ia stai puțin!", ar putea gândi unii dintre voi. „Problema asta nu a fost rezolvată deja? Waymo, automobilul

autonom al lui Google, nu a condus deja milioane de kilometri? Waymo-urile, mașinile complet autonome (sau cel puțin *aproape* complet autonome) nu circulă în prezent pe străzile din Phoenix, statul Arizona?"

Ei bine, da! A șa este! Dar nu orice kilometru de șosea este la fel. Cele mai multe distanțe sunt ușor de parcurs, poți conduce și dacă ești cu gândul aiurea. Altele sunt mult mai dificile. În momentul în care scriu această carte, automobilelor Waymo nu li se permite să meargă chiar pe oriunde: sunt „limitate geografic" la o suprafață redusă și predefinită. Tot astfel sunt și vehiculele fără șofer pe care companiile Daimler și Ford își propun să le lanseze pe șosele până în anul 2020, respectiv 2021. Sunt mașini care vin doar la comandă, ducând mai mulți pasageri care merg (aproape) în aceeași direcție, iar călătoriile le sunt limitate la o zonă de circulare predeterminată. Iar asta simplifică destul de mult problema autonomiei.

Paul Newman crede că acesta este viitorul la care ne putem aștepta din partea vehiculelor fără șofer:

> Ele vor ieși să lucreze într-o zonă pe care o cunosc foarte bine, unde posesorii lor sunt foarte convinși că vor funcționa. Prin urmare, ar putea fi o anumită parte a orașului, nu în mijlocul unei zone cu șosele mai neobișnuite sau unde s-ar putea să le iasă niște vaci în cale. Poate vor funcționa în anumite momente ale zilei sau în anumite condiții atmosferice. Vor fi exploatate ca un serviciu de transport în comun.

Nu este chiar același lucru cu autonomia deplină. Iată, în cele ce urmează, opinia lui Jack Stilgoe despre compromisul necesar:

Ceea ce seamănă cu un sistem autonom este, de fapt, un sistem pe care lumea este constrânsă să-l facă să pară autonom.

Viziunea în care am ajuns să credem este ca o iluzie optică. Un miraj care promite o limuzină cu un șofer particular stilat pentru toată lumea, însă care, dacă o cercetăm mai bine, nu va fi decât un microbuz local.

Dacă încă mai aveți nevoie de argumente, îi voi lăsa ultimul cuvânt uneia dintre cele mai importante reviste auto din America, și anume *Car and Driver*:

> Nicio companie producătoare de mașini nu se așteaptă ca utopia futuristă a unor străzi lipsite de accidente și înțesate de vehicule fără șofer să prindă contur în viitorul apropiat și nici mai devreme de câteva decenii. Companiile vor, însă, să fie luate în serios de bursa de pe Wall Street, dar și să înflăcăreze imaginația unui public din ce în ce mai puțin interesat de condusul unui automobil. Iar, între timp, speră să vândă o mulțime de autovehicule dotate cu cea mai recentă și mai sofisticată tehnologie de asistare la conducere (*driver-assistance technology*).[39]

Dar ce este cu tehnologia asta de asistare la conducere? În definitiv, mașinile fără șofer nu sunt o propunere de genul totul sau nimic.

În categoria tehnologiei „fără șofer" există șase niveluri diferite: de la nivelul 0 – fără niciun fel de automatizare și până la nivelul 5 – visul autonomiei complete. Între acestea, ea variază de la *cruise control* (sau regulator de viteză) (nivelul 2), la vehiculele autonome cu delimitare zonală (nivelul 4), iar în limbajul curent sunt cunoscute drept: nivelul 1 (*feet off*), nivelul 2 (*hands off*), nivelul 3 (*eyes off*) și nivelul 4 (*brain off*).*

* În conformitate cu definițiile propuse de Society of Automotive Engineers: nivelul 1 (*feet off, hands on*), șoferul împarte controlul vehiculului cu algoritmul (ca în cazul parcării asistate sau al monitorizării automate a vitezei); nivelul 2 (*hands off*) presupune că sistemul automat preia pe deplin controlul vehiculului, accelerând, punând frână și răsucind volanul, în vreme ce șoferul trebuie să intervină doar în cazul în care sistemul nu face față; nivelul 3 (*eyes off*) înseamnă că automobilul se poate descurca singur chiar și în cazurile în care ar fi nevoie să frâneze brusc, ceea ce înseamnă că șoferul uman își poate îndrepta atenția spre alte sarcini (cum ar fi să scrie un SMS), deși trebuie să fie pregătit să intervină în cazuri extreme. Nivelul 4 de automatizare (*brain/mind off*) seamănă cu nivelul 3, numai că șoferul mai trebuie să intervină doar în cazuri cu totul speciale (într-un ambuteiaj etc.), vehiculul fiind capabil să ia chiar și decizii legate de

Așa că nivelul 5 poate că nu este chiar în orizontul nostru imediat de așteptare, iar nivelul 4 nu va fi chiar ceea ce pare, însă ne mai așteaptă foarte multă automatizare în timp ce progresăm la nivelul superior. Oare chiar ne-ar deranja să trecem încet-încet de la un nivel la altul cu automobilele noastre personale? Și să construim mașini cu volane, pedale de frână, șoferi adevărați pe scaunul șoferului, lăsând ființa umană să intervină și să preia comanda doar în situațiile de urgență? Chiar nu ne vom mulțumi cu aceste progrese până când va avansa suficient tehnologia?

Din nefericire, lucrurile nu sunt chiar atât de simple. Pentru că povestea aceasta mai are o întorsătură neașteptată. O pleiadă întreagă de alte probleme, de fapt. Un obstacol inevitabil pentru orice altceva în afară de conducerea complet lipsită de elementul uman.

MEZINUL COMPANIEI

Printre piloții Air France, Pierre-Cédric Bonin era cunoscut drept „mezinul companiei".[40] El s-a alăturat Air France la frageda vârstă de 26 de ani, având la activ doar câteva sute de ore de zbor, și a crescut profesional printre flota de Airbus-uri a companiei. Până când a urcat la bordul fatidicului AF447, la vârsta de 32 de ani, reușise să acumuleze respectabilul număr de 2 936 de ore de zbor, deși asta îl făcea să fie totuși cel mai puțin experimentat dintre cei trei piloți aflați în aeronavă.[41]

Cu toate acestea, Bonin a fost cel care s-a așezat la comanda zborului 447 al Air France pe 31 mai 2009, a scos avionul de pe pista Aeroportului Internațional Galeão din Rio de Janeiro și s-a îndreptat spre casă, spre Paris.[42]

întreruperea călătorie. Nivelul 5 (*steering wheel optional*) aparține automobilelor (precum un taxi robotizat) care nu mai necesită nicio intervenție umană (n. red.).

Era un Airbus A330, unul dintre cele mai sofisticate avioane comerciale construite vreodată. Sistemul său autopilot era atât de modern, încât putea, practic, să efectueze un întreg zbor neasistat, cu excepția manevrelor de decolare și aterizare. Chiar și atunci când pilotul era la comandă, avionul avea încorporate diverse mecanisme de siguranță pentru a reduce la minimum riscul de apariție al erorilor umane.

Numai că întotdeauna există un pericol nebănuit care apare dacă realizezi un sistem automat care poate gestiona în siguranță aproape orice problemă pe care au reușit s-o anticipeze creatorii săi: dacă se așteaptă de la un pilot să preia manevrarea avionului numai în împrejurări excepționale, acesta nu-și va mai păstra aptitudinile necesare pentru a opera singur sistemul, pentru a conduce neajutat aeronava. Așa că omul va ajunge să aibă doar o experiență foarte redusă la care să se raporteze atunci când va întâlni o situație de urgență pe care nimeni n-a prevăzut-o.

Și exact asta s-a întâmplat cu zborul Air France 447: cu toate că Bonin acumulase mii de ore de antrenament în carlinga unui Airbus, experiența lui reală de pilotare manuală a unui A330 era minimă. Rolul său ca pilot fusese, în cea mai mare parte, acela de a monitoriza sistemul automat. Asta însemna că atunci când pilotul automat s-a deconectat, în timpul acelui zbor de noapte, Bonin nu a știut cum să piloteze avionul în siguranță.[43]

Dificultățile au început când s-au format cristale de gheață în interiorul senzorilor din fuzelaj care măsurau viteza vântului. Incapabil să facă o măsurătoare precisă, pilotul automat a declanșat o alarmă în cabina piloților și a transferat responsabilitatea echipajului uman. Aceasta nu era o chestiune îngrijorătoare în sine. Însă atunci când avionul a ajuns într-o zonă restrânsă cu turbulențe, Bonin cel lipsit de experiență a reacționat exagerat: când avionul a început să se răsucească ușor pe partea dreaptă, Bonin a

apucat manșa și a tras-o spre stânga. În același moment, ceea ce este foarte important, a tras manșa înapoi, ceea ce a făcut ca avionul să urce brusc și abrupt.[44]

Pe când aerul din jurul avionului se rarefia, Bonin continua să tragă cu putere de manșă înapoi, până ce botul avionului a ajuns atât de înălțat, încât aerul nu mai putea să circule lin de-a lungul aripilor, așa că acestea s-au transformat în adevărate paravânturi, iar aeronava și-a pierdut brusc portanța, intrând în cădere liberă cu botul îndreptat în sus.

Alarmele sunau în carlingă. Căpitanul a năvălit înăuntru din cabina de odihnă. AF447 se prăbușea în ocean cu o viteză de 3 000 de metri pe minut.

De acum, cristalele de gheață se topiseră, nu mai exista nicio defecțiune mecanică, iar oceanul era destul de departe sub ei pentru ca piloții să poată redresa avionul la timp. Bonin și copilotul său ar fi putut să-i salveze cu ușurință pe toți cei de la bord, în numai 10–15 secunde, prin simpla împingere a manșei înainte, aplecând botul avionului în jos și lăsând aerul să circule din nou lin peste aripile aparatului de zbor.[45]

Din cauza panicii care îl cuprinsese, Bonin lăsase însă manșa trasă înapoi. Nimeni nu și-a dat seama că el provocase toată drama. S-au scurs secunde prețioase. Căpitanul a sugerat să echilibreze aripile avionului. S-au întrebat, repede, dacă avionul cobora sau urca. Atunci, la 2 400 de metri distanță de nivelul mării, copilotul a preluat controlul avionului.[46]

„Urcă... urcă... urcă... urcă...", se aude copilotul strigând.

„Dar am avut tot timpul manșa trasă înapoi!" răspunde Bonin.

Atunci căpitanului i-a picat fisa. Și-a dat, în sfârșit, seama că de mai mult de trei minute erau în cădere liberă, în stare de blocaj aerodinamic și le-a ordonat să coboare botul avionului. Prea târziu, însă. În mod tragic, ajunseseră prea aproape de suprafața apei. Bonin strigă: „Fir-ar să fie! Ne vom prăbuși. Nu se poate!"[47] Câteva clipe mai târziu, aeronava a plonjat în Oceanul Atlantic, omorându-i pe toți cei 228 de oameni de la bord.

Cu 26 de ani înainte de prăbușirea avionului Air France, în 1983, psihologul Lisanne Bainbridge a scris o lucrare fundamentală despre pericolele care ne pândesc atunci când ajungem să ne bazăm prea mult pe sistemele automate.[48] Construiește o mașină care îmbunătățește performanțele umane, arăta ea, și aceasta va duce – în mod ironic – la o diminuare a aptitudinilor umane.

În prezent, suntem cu toții, într-o oarecare măsură, martorii acestui fenomen. De aceea oamenii nu-și mai pot aminti numerele de telefon, de aceea mulți dintre noi ne chinuim să ne descifrăm scrisul de mână și de aceea foarte mulți nu mai pot să ajungă nicăieri fără GPS. Dată fiind tehnologia, care face totul pentru noi, mai rămân puține prilejuri să ne exersăm abilitățile.

Există o anumită îngrijorare că același fenomen s-ar putea să se manifeste și odată cu apariția vehiculelor care se vor conduce singure – unde mizele vor fi cu mult mai mari decât în cazul scrisului de mână. Până când vom ajunge la autonomia deplină, mașina va mai preda uneori controlul, pe neașteptate, șoferului. Vom reuși să ne aducem instinctiv aminte ce avem de făcut? Iar șoferii adolescenți din viitor vor avea vreodată șansa să ajungă să stăpânească toate aptitudinile necesare?

Chiar dacă toți șoferii reușesc să rămână competenți*(și interpretăm cu multă generozitate verbul „să rămână"), mai există o problemă cu care va trebui să ne confruntăm. Pentru că tot ce i se cere șoferului uman să facă înainte de întreruperea pilotului automat este, de asemenea, important. Există

* Există diverse măsuri care se pot lua pentru a remedia problemele care apar din cauză că nu ai avut prea mult timp să exersezi conducerea manuală a unui avion. De exemplu, de la prăbușirea avionului Air France se pune acum accentul pe antrenarea noilor piloți, astfel încât să fie capabili să piloteze avionul atunci când pilotul automat nu reușește să o facă. De asemenea, ei sunt stimulați ca, la intervale regulate, să oprească sistemul automat pentru a-și păstra îndemânarea de a pilota manual avionul (n. a.).

numai două posibilități. Și – așa cum subliniază Bainbridge – niciuna nu este foarte atrăgătoare.

O mașină de nivelul 2, *hands-off*, se va aștepta ca șoferul să fie foarte atent la drum în orice moment.[49] Nu este suficient de performantă, încât să fie lăsată de una singură și va avea nevoie de o supraveghere atentă. Revista *Wired* descria la un moment dat acest nivel astfel: e „ca atunci când îi ceri unui copilaș care abia a învățat să meargă să te ajute să speli vasele."[50]

În momentul scrierii cărții de față, sistemul autopilot al unei mașini Tesla este un exemplu în acest sens.[51] În prezent, este ca o funcție mai sofisticată de *cruise control* – va conduce, va frâna și va accelera pe autostradă, dar se așteaptă ca șoferul să fie vigilent și atent, pregătit să intervină în orice moment. Ca să se asigure că ești atent, ea va declanșa o alarmă dacă îți iei mâinile de pe volan prea mult timp.

Dar, așa cum a scris Bainbridge în lucrarea ei, această abordare va ajunge să provoace accidente, pentru că este, pur și simplu, lipsit de realism să te aștepți ca oamenii să fie vigilenți:

> Este imposibil, chiar și pentru o persoană foarte motivată, să-și păstreze atenția vizuală îndreptată în mod eficient spre o sursă de informații în care se întâmplă foarte puține lucruri, pentru mai mult de aproximativ o jumătate de oră.[52]

Sunt unele dovezi că oamenii au întâmpinat dificultăți în a ține cont de insistența companiei Tesla ca ei să fie atenți la drum. Joshua Brown, care a murit la volanul mașinii sale Tesla în anul 2016, utiliza funcția de autopilot de 37 de minute și 30 de secunde când mașina a lovit un camion care traversa banda pe care conducea el. Din cercetările efectuate de Administrația pentru Siguranța Traficului Rutier a rezultat că Brown nu se uita la drum în momentul impactului.[53] Accidentul a ajuns pe prima pagină a ziarelor din toată lumea, însă asta nu i-a oprit pe unii nesăbuiți de pe YouTube să posteze materiale

video în care se arăta, cu mult entuziasm, cum să-ți păcălești mașina, astfel încât aceasta să creadă că ești atent la drum. Chipurile, dacă lipești o doză de Red Bull de volan [54] sau îndeși o portocală în volan[55], asta va împiedica mașina să declanșeze alarmele acelea sâcâitoare care îți amintesc de responsabilitățile pe care le ai.

Alte programe întâmpină aceleași probleme. Cu toate că mașinile Uber fără șofer au nevoie de intervenție umană la fiecare 20 de kilometri[56], a-i face pe șoferi să rămână atenți la drum este un adevărat calvar. Pe 18 martie 2018, un vehicul Uber autonom a lovit mortal un pieton. Înregistrarea video realizată din interiorul mașinii a arătat că „monitorul uman" care stătea în spatele volanului nu se uita la drum în minutele imediat premergătoare coliziunii.[57]

Aceasta este o problemă serioasă, însă există și o soluție alternativă. Companiile producătoare de mașini ar putea să accepte ideea că oamenii sunt oameni și să recunoască faptul că atenția ne poate fi distrasă. În definitiv, să poți să citești o carte în timp ce conduci face parte din ceea ce ne atrage, ne *seduce*, la mașinile care se conduc singure. Este diferența principală dintre nivelul 2 (*hands off*) și nivelul 3 (*eyes off*).

Cel din urmă presupune o provocare tehnică superioară față de nivelul 2, însă unii producători auto au început deja să-și construiască mașinile ținând cont de lipsa noastră de atenție. Tehnologia Audi *traffic-jam pilot* este un asemenea exemplu.[58] Aceasta poate să preia complet controlul autovehiculului atunci când rulezi în condiții de trafic cu viteză redusă pe șosea, permițându-ți să te relaxezi și să te bucuri de călătorie. Trebuie doar să fii pregătit să intervii dacă apare ceva neprevăzut.*

* Fiind cu un pas înainte față de automatizarea parțială, vehiculele de nivel 3, cum este Audi dotat cu *traffic-jam pilot*, pot să preia controlul în unele situații, dacă sunt îndeplinite anumite condiții. Șoferul încă mai trebuie să fie pregătit să intervină atunci când automobilul se confruntă cu un scenariu pe care nu îl înțelege,

Există un motiv pentru care Audi şi-a limitat sistemul la traficul cu viteză redusă de pe drumurile cu acces limitat. Riscul unor catastrofe este mai mic într-un trafic aglomerat. Iar acesta este un lucru important. Pentru că, în momentul în care omul încetează să mai supravegheze drumul, se poate genera cea mai proastă combinaţie de împrejurări posibilă atunci când apare o situaţie de urgenţă.

Un şofer care nu este atent va dispune de foarte puţin timp ca să evalueze situaţia din jur şi să decidă ce are de făcut. Imaginaţi-vă că vă aflaţi într-o maşină care se conduce singură, auziţi un semnal de alarmă şi ridicaţi ochii din carte tocmai la timp ca să vedeţi în faţă un camion care-şi pierde încărcătura pe drumul pe care sunteţi. Într-o fracţiune de secundă trebuie să procesaţi informaţia din jur: motoreta de pe banda stângă, furgoneta din faţă care frânează brusc, maşina din unghiul mort din partea dreaptă. Vă veţi trezi că nu sunteţi deloc familiarizat cu drumul chiar în momentul în care ar trebui să-l cunoaşteţi cel mai bine; adăugaţi la asta lipsa de experienţă practică şi veţi fi cât se poate de slab pregătit pentru a face faţă situaţiilor care cer cel mai ridicat nivel de îndemânare.

Acest fapt a fost confirmat şi în simulările cu maşini fără şofer. Un studiu în care li se permitea participanţilor să citească o carte sau să se joace pe telefonul mobil în timp ce maşina se conducea singură a descoperit că a fost nevoie de până la 40 de secunde după declanşarea unei alarme pentru ca oamenii să recâştige controlul corespunzător asupra vehiculului.[59] Este exact ce s-a întâmplat şi cu zborul 447 al Air France. Căpitanului Dubois, care ar fi trebuit să poată să salveze cu uşurinţă avionul de la prăbuşire, a avut nevoie de aproape un minut ca să-şi dea seama ce se întâmpla şi să găsească soluţia simplă care ar fi rezolvat problema.[60]

dar nu mai este obligat să monitorizeze permanent drumul şi vehiculul. Nivelul acesta seamănă mai mult cu situaţia în care trebuie să convingi un adolescent să spele vasele (n. a.).

În mod ironic, cu cât tehnologia de conducere fără șofer se îmbunătățește, cu atât mai grave devin problemele care pot să apară. Un sistem de conducere pe pilot automat neglijent conceput, care declanșează alarma la fiecare 15 minute, va menține șoferul permanent vigilent și îl va obliga să practice șofatul în mod constant. Cele la care trebuie să fii într-adevăr atent sunt sistemele automate performante și sofisticate, care sunt *aproape* întotdeauna demne de încredere.

De aceea Gill Pratt, care conduce institutul de cercetare al Toyota, a spus următoarele:

> Cel mai rău scenariu este o mașină care să aibă nevoie de intervenția șoferului numai o dată la 300 000 de kilometri... Un om obișnuit, care are o mașină [nouă] la fiecare 160 000 de kilometri, nu va remarca niciodată [când sistemul automat îi va ceda controlul șoferului]. Însă, din când în când, poate o dată la fiecare două mașini pe care le voi deține, va exista acea singură ocazie când se va auzi brusc: „bip, bip, bip, acum este rândul tău!" Iar persoana în cauză, care nu s-a mai întâlnit cu această situație, probabil, de ani de zile [...], nu va fi deloc pregătită să-i facă față.[61]

MARILE SPERANȚE

În ciuda tuturor acestor probleme, există motive întemeiate pentru a ne dori să progresăm către un viitor cu vehicule care se conduc singure. Binele continuă să învingă răul. Conducerea mașinii rămâne una dintre cele mai importante cauze ale deceselor care ar putea fi prevenite din întreaga lume. Dacă tehnologia este capabilă, oricât de puțin, să reducă numărul deceselor pe șosele, ai putea susține că ar fi lipsit de etică să *nu* o pui în practică.

Și mai există avantaje, și nu puține: chiar și sistemele simple de asistare la conducere pot contribui la reducerea

consumului de combustibil[62] și la decongestionarea traficului.[63] În plus – haideți să fim sinceri! – ideea de a-ți lua mâinile de pe volan fie și numai pentru o clipă, atunci când conduci cu 100 km/h, e de-a dreptul... cool!

Revenind, însă, la avertismentele lui Bainbridge, acestea fac aluzie la o problemă care are de-a face cu modul în care este concepută tehnologia de conducere autonomă a vehiculelor în prezent.

Să luăm Tesla, de exemplu, unul dintre primii fabricanți de mașini care au scos pe piață un model autonom. Nu există nicio îndoială că sistemul lor a avut un impact pozitiv, făcând conducerea mașinii mai sigură pentru cei care o folosesc – nu trebuie să cauți prea mult ca să găsești on-line materiale video cu sistemul „Forward Collision Warning", care recunoaște riscul producerii unui accident chiar înaintea șoferului, declanșând o alarmă și salvând mașina de la coliziune.[64]

Numai că apare o ușoară neconcordanță între ceea ce pot să facă mașinile – în esență, cu niște senzori pretențioși pentru parcarea laterală și un sistem inteligent de *cruise control* – și limbajul utilizat pentru descrierea acestora. În octombrie 2016, de exemplu, compania a anunțat că „toate automobilele Tesla produse în prezent sunt dotate cu hardware *full self-driving*"*. După un articol din *The Verge*, Elon Musk, arhitectul companiei Tesla, a adăugat: „De acum înainte, actualizarea autonomiei complete va constitui o procedură standard la toate autovehiculele Tesla."[65] Iar conceptul de „autonomie completă" vine, fără îndoială, în

* În februarie 2018, când scriu această carte, hardware-ul *full self-driving* este un pachet opțional care poate fi achitat la cumpărare, cu toate că vehiculul încă nu deține software-ul care să permită deplasări cu o conducere complet automată. Pe site-ul Tesla se precizează: „Nu putem ști cu precizie când va deveni disponibil fiecare element funcțional descris mai sus." Vezi https://www.tesla.com/en_GB/blog/all-tesla-cars-being-produced-now-have-full-self-driving-hardware (link accesat cel mai recent la 18 martie 2019) (n. a.).

conflict cu avertizarea pe care conducătorii auto trebuie să o accepte înainte de a folosi funcția autopilot actuală: „Trebuie să mențineți controlul asupra vehiculului dumneavoastră și să fiți responsabil de acțiunile acestuia."[66]

Speranțele își au importanța lor. Este posibil să nu fiți de acord, însă eu cred că o consecință inevitabilă a folosirii unui limbaj ce induce în eroare de către reprezentanții unui brand de încredere este tocmai apariția acelor persoane care își îndeasă portocale în volan – sau, mai rău, așa cum am descoperit prin cotloanele mai întunecate ale web-ului, care concep și vând dispozitive care „[permit] adoptanților timpurii* să [conducă în timp ce] opresc avertizările de verificare ale autopilotului sau le reduc volumul"**.

Desigur că din toată industria de automobile, Tesla nu este singura vinovată. Orice companie din lumea asta se folosește de fanteziile noastre ca să-și vândă produsele. În ceea ce mă privește, însă, există o diferență între a cumpăra un parfum pentru că sunt de părere că mă va face mai atrăgătoare și a cumpăra o mașină deoarece consider că sistemul ei complet autonom mă va menține în siguranță.

Lăsând la o parte strategiile de marketing, nu pot decât să mă întreb dacă nu cumva am avut tot timpul o impresie greșită despre vehiculele fără șofer.

Știm de acum că oamenii se pricep să înțeleagă subtilitățile, să analizeze contextul, să se bazeze pe experiența acumulată

* Un adoptant timpuriu sau adoptant precoce este unul dintre primii clienți ai unei companii, unul dintre primii consumatori ai unui produs sau unul dintre primii utilizatori ai unei tehnologii (cf. articolul din Wikipedia) (n. tr.).

** Este un produs care există în realitate, se numește *The Autopilot Buddy* și îl poți cumpăra la prețul „de nimic" de 179 de dolari. Este demn de menționat că pe website se precizează, cu litere mici: „*Autopilot Buddy* nu trebuie utilizat niciodată pe străzile publice." Vezi https://www.autopilotbuddy.com/ (link accesat cel mai recent la 19 martie 2019) (n. a.).

și să observe tipare, pattern-uri. Nu ne pricepem mai deloc să fim atenți, precișI, consecvenți și pe deplin conștienți de ceea ce ne înconjoară. Deținem, pe scurt, abilități și capacități 100% opuse față de acelea ale algoritmilor.

De ce nu am urma, așadar, exemplul software-ului din medicină de depistare a tumorilor, lăsând abilitățile mașinii să vină în completarea celor umane și făcând, astfel, ca abilităților amândurora să se dezvolte? Până când vom ajunge la autonomia completă, de ce nu am răsturna ecuația cu susul în jos și nu am tinde spre un sistem autonom care să vină în sprijinul conducătorului auto, iar nu viceversa? De ce nu am folosi o plasă de siguranță, cum este ABS-ul sau sistemul de control al tracțiunii, care monitorizează cu răbdare drumul și rămâne în alertă, pândind vreun pericol pe care șoferul l-a trecut cu vederea. Prin urmare, să folosim sistemul computerizat nu ca pe un conducător auto, ci ca pe un supraveghetor.

Aceasta este ideea care stă la baza muncii depuse de institutul de cercetări al Toyota. Cei de acolo includ în construcția mașinii două sisteme. Sistemul *chauffeur**, care – asemănător tehnologiei *traffic-jam pilot* de la Audi – ar putea prelua controlul într-un trafic foarte aglomerat, și sistemul *guardian***, care rulează în fundal atunci când conduce șoferul uman și acționează ca o plasă de protecție[67], reducând riscul unui accident dacă se ivește ceva ce șoferul nu a observat.

Compania Volvo a adoptat o strategie similară. Sistemului său de frânare automată în caz de urgență***, care încetinește automat automobilul dacă se apropie prea mult de acela din fața sa, i se recunosc pe scară largă meritele pentru impresionantul record de siguranță al lui Volvo XC90. Din anul 2002, când mașina a apărut pentru prima oară pe piața din Marea Britanie,

* „Șofer" (fr. în original) (n. tr.).

** „Păzitor, supraveghetor" (engl.) (n. tr.).

*** Autonomous Emergency Breaking (AEB) (engl.), „sistem de frânare autonom" (n. tr.).

au fost achiziționate peste 50 000 de vehicule și niciun șofer sau pasager din interiorul lor nu a fost ucis în vreo coliziune.[68]

Ca și în cazul tehnologiei „fără șofer", despre care se discută cu atâta înflăcărare, va trebui să așteptăm și să vedem cum se vor termina toate acestea. Însă un lucru este sigur: cu timpul, conducerea autonomă ne va învăța câteva lecții care vor putea fi aplicate și în afara universului automobilistic. Nu numai despre cum să transferăm controlul, ci și cum să dăm dovadă de realism când ne formulăm așteptările despre ceea ce pot algoritmii să facă.

Pentru ca acest demers să fie viabil, trebuie să ne schimbăm felul de a gândi și să renunțăm la ideea că automobilele trebuie să funcționeze perfect de fiecare dată și să acceptăm că, deși defecțiunile mecanice ale vehiculelor vor deveni un eveniment rar întâlnit, mai avem drum lung de străbătut până când „defectarea" algoritmilor să fie și ea la fel de rară.

Așadar, știind că greșelile sunt inevitabile, știind că, în cazul în care vom merge mai departe, nu avem altă alternativă decât să acceptăm incertitudinea, dilemele morale din lumea mașinilor fără șofer ne vor determina să decidem cât de bună trebuie să fie o astfel de mașină înainte de a fi dispuși să o lăsăm să circule liberă pe străzile noastre. Este o întrebare importantă, iar ea rămâne valabilă și într-un context diferit. Cât de bun înseamnă „destul de bun"? Odată ce ai construit un algoritm defectuos care *poate* să calculeze ceva, ar trebui să-l și lași să o facă?

Capitolul 6

INFRACȚIONALITATEA

Era o zi călduroasă de iulie din 1995, când o studentă de 22 de ani și-a strâns cărțile, a părăsit biblioteca din Leeds și s-a îndreptat spre mașina ei. Își petrecuse ziua punând la punct ultimele detalii ale lucrării de diplomă, iar acum era liberă să se bucure de restul verii. Tocmai când se așeza în mașină la volan, a auzit pe cineva alergând în direcția ei prin parcarea supraetajată. Înainte de a putea să reacționeze, un bărbat s-a aplecat spre ea prin geamul deschis al mașinii și i-a pus cuțitul la gât. A silit-o să se așeze pe bancheta din spate, a legat-o, i-a lipit pleoapele cu lipici, a trecut la volan și a plecat cu mașina.

După o călătorie îngrozitoare, bărbatul a oprit mașina pe un taluz năpădit de buruieni. Fata a auzit un zgomot când necunoscutul și-a rabatat scaunul, iar după aceea un foșnet când a început să-și desfacă șlițul. Știa că intenționa să o violeze. Luptându-se orbește, fata și-a ridicat genunchii la piept și apoi i-a împins în față cu toată puterea, forțându-l să se dea înapoi. În timp ce ea îl lovea cu picioarele și se zbătea, bărbatul s-a tăiat la degete cu cuțitul pe care-l ținea în mână, iar sângele i-a picurat pe scaunele mașinii. A lovit-o pe tânără de două ori peste față, apoi, spre imensa ei ușurare, bărbatul a ieșit din mașină și a plecat. La două ore de la începerea calvarului, studenta a fost găsită rătăcind pe Globe Road din Leeds, ametită și cu părul despletit, cu cămașa sfâșiată, cu fața înroșită acolo unde bărbatul o lovise și cu pleoapele lipite cu lipici.[1]

Agresiunile sexuale asupra persoanelor necunoscute, ca în acest caz, sunt extrem de rare, însă, atunci când au loc, ele par să facă parte dintr-o serie. Mai mult ca sigur, nu era prima oară când bărbatul acela comitea o asemenea faptă. Când poliția a analizat sângele din mașină, a descoperit că ADN-ul bărbatului era identic cu acela dintr-o probă prelevată de la locul unde fusese comis un viol cu doi ani în urmă, într-o altă parcare supraetajată. Acel atac avusese loc în Nottingham, cu vreo sută de kilometri mai spre sud. Iar, ca urmare a unui apel lansat în programul *Crimewatch* al BBC, poliția a reușit să lege cazul de alte trei incidente, petrecute cu zece ani înainte, în Bradford, Leeds și Leicester.[2]

Numai că găsirea acestui violator în serie nu avea să fie ușoară. Toate infracțiunile fuseseră comise pe o arie de 7 046 km² – o suprafață imensă din Marea Britanie. În plus, ele le ofereau polițiștilor un număr uluitor de mare de potențiali suspecți – 33 628 în total – fiecare din ei trebuind să fie eliminat din cercetările poliției sau investigat.[3]

Se impunea, așadar, organizarea unei operațiuni de căutare de proporții uriașe, ceea ce nu se întâmpla pentru prima oară, căci, după agresiunile din urmă cu zece ani, fusese întreprinsă o razie de proporții, numai că, în ciuda faptului că polițiștii bătuseră la 14 153 de uși și recoltaseră numeroase probe biologice (fire de păr, bețișoare cu probe de salivă pentru prelevarea ADN-ului și altele), ancheta nu dusese, în cele din urmă, nicăieri. Era un risc foarte mare ca și această nouă căutare să aibă aceeași soartă, până când Kim Rossmo, un fost polițist canadian, și algoritmul pe care tocmai îl crease au fost chemați în ajutor.[4]

Rossmo avusese o idee îndrăzneață: în loc să țină cont de imensul volum de dovezi deja existente, algoritmul lui ignora, practic, totul. În schimb, își concentra atenția exclusiv asupra unui singur factor: localizarea geografică.

Rossmo considera că, poate, un făptaș nu alege la întâmplare *locul* unde își atacă victima. Poate că selectarea acestuia nu este o decizie luată absolut liber sau conștient. Chiar dacă agresiunile avuseseră loc de-a lungul și de-a latul Angliei, Rossmo s-a întrebat dacă nu cumva ar putea exista un tipar involuntar ascuns în repartizarea geografică a infracțiunilor – un tipar suficient de simplu, încât să poată fi exploatat. Era o șansă, credea el, ca locul în care aveau loc infracțiunile să îl trădeze pe cel de unde provenea făptașul. Cazul violatorului în serie era un prilej de a-și pune teoria la încercare.

OPERAȚIUNEA LYNX ȘI ASPERSORUL

Rossmo nu era prima persoană care sugera că infractorii alcătuiesc, fără să-și dea seama, tipare geografice. Ideile lui au o istorie care ajunge până prin anii 1820, când André-Michel Guerry, un avocat devenit statistician care lucra la Ministerul de Justiție al Franței, a început să strângă dosarele despre violurile, crimele și furturile care aveau loc în diverse regiuni ale țării.[5]

Deși colectarea acestui tip de statistici pare un lucru destul de obișnuit în zilele noastre, în perioada aceea, matematica și statistica fuseseră aplicate doar în cadrul științelor exacte, în care ecuațiile erau folosite pentru a descrie în mod elegant legile fizice ale universului, pentru trasarea unor lucruri precum trasarea traiectoriei unei planete pe cer sau calcularea forțelor care acționează într-un motor cu aburi. Până atunci nimeni nu-și dăduse osteneala să strângă informații despre infracțiuni. Nimeni nu avea habar ce-ar fi trebuit să calculeze, cum să calculeze sau cu ce frecvență ar trebui să facă respectivele calcule. De altfel – gândeau oamenii pe vremea aceea – ce rost ar fi avut? Un bărbat era puternic, independent din fire și se deplasa de colo-colo, acționând după propria

voință. Purtarea sa era imposibil de surprins de neînsemnata disciplină a statisticii.[6]

Însă analiza făcută de Guerry adevăratului său recensământ al infractorilor de pe teritoriul Franței sugera cu totul altceva. Oriunde te-ai fi aflat în Franța, a descoperit el, infracțiunile comise prezentau anumite tipare care puteau fi recunoscute: tipare despre tipul infracțiunilor, modul în care fuseseră comise și de către cine. Persoanele tinere comiteau mai multe infracțiuni decât cele în vârstă, bărbații mai mult decât femeile, săracii mai mult decât bogații. În mod curios, a devenit în curând limpede că aceste tipare nu se modificau odată cu trecerea timpului. Fiecare regiune din Franța avea propria statistică pe categorii de infracțiuni, care abia dacă înregistra vreo schimbare de la an la an. Cu o precizie aproape înspăimântătoare, numărul furturilor, violurilor și omorurilor se repeta de la un an la altul. Până și metodele folosite de ucigași erau predictibile. Asta însemna că Guerry și colegii săi puteau să aleagă o zonă și să spună în avans, cu precizie, câte omucideri cu cuțitul, sabia, prin lapidare, strangulare sau înec se puteau aștepta să aibă loc într-un anumit an.[7]

Așadar, până la urmă, poate că nu era o chestiune de liber arbitru al criminalului. Crima nu se produce la întâmplare; oamenii sunt previzibili. La aproape două secole de la descoperirea ei de către Guerry, Kim Rossmo voia să exploateze exact această predictibilitate.

Munca lui Guerry se concentrase asupra tiparelor pe care le găsise la nivel național și regional, dar s-a dovedit că și la nivel individual persoanele care comiteau infracțiuni creau tipare geografice pe care te puteai baza. Asemenea nouă, infractorii au tendința să rămână în teritorii care le sunt familiare. Operează local. Asta înseamnă că până și cele mai grave infracțiuni vor fi duse la îndeplinire în apropierea locului în care își duce viața făptașul. Iar pe măsură ce te îndepărtezi de scena crimei, șansa de a găsi sălașul acestuia se reduce

treptat[8], un efect cunoscut de criminologi sub denumirea de „deteriorarea relațiilor odată cu creșterea distanței"*.

Pe de altă parte, este puțin probabil ca niște criminalii în serie să vizeze victime care locuiesc *foarte* aproape de ei, pentru a evita să atragă în mod inutil atenția poliției asupra lor sau să riște să fie recunoscuți de vecini. De aici rezultă apariția a ceea ce se numește o „zonă tampon", care include locuința infractorilor, o zonă în care sunt șanse foarte mici ca aceștia să comită vreo infracțiune.[9]

Aceste două tipare de bază – deteriorarea relațiilor odată cu creșterea distanței și zona tampon –, ascunse în localizarea geografică a celor mai grave crime, au stat la baza algoritmului lui Rossmo. Începând de la locul crimei, prins cu un ac pe o hartă, Rossmo și-a dat seama că putea să echilibreze matematic acești doi factori și să schițeze o imagine a ariei probabile în care era posibil să locuiască infractorul.

Imaginea aceasta nu este deosebit de utilă atunci când a fost comisă o singură infracțiune. În lipsa unor informații suficiente pentru a merge mai departe, așa-numitul „algoritm de geo-profilare a infractorului"** nu-ți va spune cu mult mai mult decât minunatul și demodatul bun-simț. Dar pe măsură ce se adaugă tot mai multe infracțiuni, imaginea începe să se contureze, aducând treptat la lumină o hartă a orașului care marchează zonele în care este cel mai probabil ca făptașul să fie prins.

Este ca și când infractorul ar fi un aspersor care se rotește. Așa cum ar fi greu să prevezi exact unde va cădea următoarea picătură de apă, nu poți să anticipezi nici unde urmează să atace acesta. Însă, după ce apa a stropit un timp, iar pe jos au căzut numeroși stropi, este destul de ușor să-ți dai seama,

* *Distance decay* (engl.), vezi glosarul, p. 317 (n. red.).

** Profilarea infractorului (*criminal profiling*, engl.) reprezintă procesul realizat de specialiștii în psihologie judiciară pentru a încerca să realizeze profilul psihologic al unui răufăcător, caracteristicile personalității acestuia (n. red.).

din tiparul pe care-l alcătuiesc picăturile, unde este așezat, cu aproximație, aspersorul.

Exact asta s-a întâmplat și cu algoritmul lui Rossmo în Operațiunea Lynx – goana după violatorul în serie. Echipa știa acum unde avuseseră loc cele cinci crime și mai cunoștea câteva locuri unde atacatorul folosise un card de credit furat pentru a-și cumpăra băuturi alcoolice, țigări și un joc video. Doar în baza acestor localizări, algoritmul a scos în evidență două zone principale în care considera că era posibil să locuiască făptuitorul: Millgarth și Killingbeck, ambele situate în suburbiile orașului Leeds.[10]

Între timp, în centrul de coordonare, poliția mai dispunea de o probă importantă pentru a duce mai departe acest caz: o amprentă parțială lăsată de atacator la locul comiterii unei fapte anterioare. Era un eșantion prea mic pentru ca sistemul automat de recunoaștere a amprentelor să poată să răsfoiască baza de date a infractorilor condamnați în căutarea unei potriviri, așa că orice comparație trebuia efectuată meticulos, cu lupa, de către un expert care să examineze cu migală fiecare suspect pe rând. Operațiunea dura, de acum, de aproape trei ani – în pofida eforturilor susținute a 180 de ofițeri din cinci departamente diferite – și începea să-și cam epuizeze energia. Fiecare indiciu ducea, de fiecare dată, pe o linie moartă.

Ofițerii de poliție au decis să verifice manual toate amprentele găsite în cele două locuri semnalate de algoritm. Au început cu Millgarth: o căutare printre fișele cu amprente păstrate în baza de date a poliției locale nu a oferit niciun rezultat. Au trecut la Killingbeck și, după 940 de ore de căutare minuțioasă printre fișele de acolo, poliția a reușit să descopere un nume: Clive Barwell.

Acesta era un bărbat de 42 de ani, căsătorit și tatăl a patru copii, care fusese arestat pentru jaf armat în intervalul de timp dintre atacuri. În acel moment, lucra ca șofer de camion și în cursul activității făcea în mod regulat curse lungi prin

toată țara; dar locuia în Killingbeck și își vizita des mama, care locuia în Millgarth, cele două zone semnalate de algoritm.[11] Amprenta parțială nu fusese de ajuns, ea singură, ca să-l identifice în mod concludent, însă un test ADN ulterior a dovedit că Barwell comisese toate acele infracțiuni odioase. Polițiștii îl găsiseră pe infractor! Barwell și-a recunoscut vinovăția la proces, în octombrie 1999. Judecătorul l-a condamnat la opt condamnări pe viață succesive.[12]

Când totul s-a terminat, Rossmo a avut prilejul să facă un bilanț al eficacității algoritmului. Acesta nu indicase niciodată propriu-zis numele lui Barwell, dar semnalase pe hartă zonele în care poliția trebuia să-și concentreze atenția. Dacă aceasta ar fi folosit algoritmul ca să-și organizeze lista de suspecți în ordinea priorității, în funcție de unde locuia fiecare – verificând amprentele și luând mostre de ADN, pe rând, de la fiecare dintre ei –, nu ar mai fi trebuit să deranjeze un număr atât de mare de oameni nevinovați. L-ar fi găsit pe Clive Barwell după ce ar fi căutat în numai 3% din locurile din întreaga zonă.[13]

Algoritmul se dovedise, cu siguranță, eficient și se vădea a avea și alte aspecte pozitive. Deoarece nu făcea altceva decât să organizeze, în ordinea priorității, o listă cu suspecți deja existentă, nu putea fi afectat de acel tip de tendențiozitate pe care l-am întâlnit la capitolul „Justiția" și nici nu avea cum să înlocuiască munca temeinică de investigare, ci doar să o eficientizeze; prin urmare, există puține șanse ca oamenii să-i acorde vreodată o încredere exagerată.

Totodată, algoritmul acesta este și foarte flexibil. După Operațiunea „Lynx", el a fost folosit de peste 350 de agenții de combatere a criminalității din întreaga lume, inclusiv de FBI-ul american și de Royal Canadian Mounted Police*, iar informațiile pe care le oferă depășesc sfera infracționalității (spre exemplu, a fost folosit pentru depistarea bălților cu apă

* Poliția Regală Canadiană Călare (engl.) (n. tr.).

stătută care le slujesc țânțarilor drept locuri de reproducere, pe baza localizării cazurilor de malarie din Egipt)[14]. Un doctorand de-al meu de la University College London utilizează, în prezent, acest algoritm în încercarea sa de a prezice amplasările fabricilor de bombe, pe baza locurilor în care sunt folosite dispozitive explozive improvizate. Iar niște matematicieni din Londra l-au folosit chiar și pentru a încerca să-l găsească pe Banksy[15], artistul solitar, pornind de la locurilor în care au fost identificate graffiti de-ale sale.

Infracțiunile pentru rezolvarea cărora geo-profilarea dă cele mai bune rezultate (violuri în serie, omucideri și atacuri violente) sunt, din fericire, rare. În realitate, marea majoritate a infracțiunilor nu presupun genul de vânătoare de oameni de care a fost nevoie în cazul lui Clive Barwell. Dacă algoritmii ar trebui să aducă o schimbare în combaterea criminalității și în afara acestor situații extreme, atunci ar avea nevoie de un tipar geografic diferit pentru a continua să dea roade. Unul care să poată fi aplicat unui oraș ca întreg și care să poată identifica pattern-urile și ritmurile unei străzi sau ale unui colț de stradă, pe care fiecare agent de poliție le cunoaște instinctiv. Din fericire, Jack Maple avea exact ceea ce trebuie.

HĂRȚI DESPRE VIITOR

În anii 1980, mulți newyorkezi ar fi stat destul de mult pe gânduri înainte de a se încumeta să meargă cu metroul, căci acesta nu era prea frumos. Toate suprafețele erau acoperite de graffiti, vagoanele duhneau a urină stătută, pe peroane erau o mulțime de persoane care luau droguri, iar furturile și tâlhăria erau la ordinea zilei. Aproximativ 20 de oameni nevinovați erau uciși în subteran în fiecare an, ceea ce făcea din metrou un spațiu la fel de periculos ca oricare altul aflat la suprafață de oriunde din lume.

Acesta era decorul în care își desfășura activitatea de ofițer de poliție Jack Maple, care fusese promovat de puțin timp la gradul de locotenent în cadrul Transit Bureau*, iar în anii petrecuți în poliție se săturase doar să reacționeze când erau comise infracțiuni, în loc să lupte pentru combaterea acestora. Din frustrare a luat naștere o idee strălucită.

Iată ce i-a declarat Maple unui jurnalist în 1999:

> Pe un perete, pe mai bine de 16 m², am trasat fiecare stație de metrou din New York și fiecare tren în parte, apoi am folosit carioci pentru a marca fiecare infracțiune violentă, tâlhărie și furt calificat care avuseseră loc acolo. Am trecut pe hartă atât cazurile rezolvate, cât și pe cele nerezolvate.[16]

Poate că nu pare cine știe ce, însă hărțile acestea, mâzgălite pe hârtie maro cu carioca, au devenit cunoscute drept „Hărțile despre viitor" și au fost considerate revoluționare la timpul lor. Nimeni nu se gândise până atunci să încerce să analizeze infracțiunile în felul acela. Dar, în momentul în care Maple s-a dat puțin înapoi, să aibă o privire de ansamblu asupra infracțiunilor din întregul oraș în același timp, și-a dat seama că privea totul dintr-o perspectivă complet diferită și, după cum explica tot el:

> Se punea întrebarea: „de ce?". Care erau cauzele responsabile de faptul că se săvârșeau o mulțime de infracțiuni într-o anumită zonă?

Problema era că, pe atunci, fiecare apel de urgență la poliție era tratat ca un incident izolat. Dacă sunai să anunți existența unui grup de dealeri de droguri agresivi care-și făceau veacul la un colț de stradă, iar aceștia se făceau nevăzuți imediat ce

* Denumirea unuia dintre cele douăzeci de birouri ale NYPD (New York Police Department, Departamentul de Poliție al New York-ului), responsabil cu siguranța celor aproximativ 5,6 milioane de pasageri care călătoresc zilnic pe cele 25 de rute ale metroului din orașul american (n. red.).

veneau polițiștii, nu se consemna nimic care să facă legătura între plângerea ta și alte apeluri de urgență primite în momentul în care gașca revenea la locul său obișnuit. Hărțile lui Maple, în schimb, îi ofereau posibilitatea să determine cu precizie locul unde infracțiunile reprezentau o problemă cronică, iar acest lucru însemna că putea să descopere și motivele acestora:

> Aici este un centru comercial? De aceea avem o mulțime de hoți de buzunare și de jafuri? Aici este o școală? De aceea avem o problemă la ora trei după-amiaza? În vecinătate se află vreo casă abandonată? De aceea se face trafic de droguri la colțul străzii?[17]

Capacitatea de a răspunde la aceste întrebări era primul pas spre soluționarea problemelor din oraș. Așa că, în 1990, când Bill Bratton, un tip deschis la minte, a devenit șeful NYPD, Mapple i-a arătat „hărțile despre viitor" pe care le concepuse. Le-au folosit amândoi pentru a încerca să facă metroul un loc mai sigur pentru toată lumea.[18]

Și Bratton a avut, la rândul său, o idee bună. Știa că persoanele care cerșeau sau urinau în subteran, precum și acelea care săreau peste sistemul de turnichete reprezentau o mare problemă în stațiile de metrou. El a decis să le ceară polițiștilor lui să își concentreze atenția asupra rezolvării acestor delicte minore înainte de a soluționa infracțiunile mult mai grave, precum jafurile și omorurile, care erau comise, de asemenea, în proporții catastrofale în subteran.

Raționamentul său avea două aspecte. În primul rând, manifestând asprime față de orice comportament antisocial chiar la locul faptei, se transmitea în cel mai puternic mod mesajul că activitatea infracțională nu era acceptată sub nicio formă, în speranța ca astfel să înceapă să transforme percepția oamenilor despre ceea ce era „normal" să se petreacă în metrou. În al doilea rând, persoanele care evitau să plătească taxa de intrare era foarte posibil să fie tocmai infractorii care aveau să comită,

ulterior, infracțiuni mai grave și, dacă erau arestate pentru că nu plătiseră taxa, nu ar mai fi avut prilejul să le ducă la îndeplinire. După cum declara Bratton ziarului *Newsday* în 1991:

> Prinzându-i pentru că nu plăteau taxa de acces, am reușit să oprim la turnichete infractori periculoși și înarmați înainte ca aceștia să ajungă în metrou și să facă ravagii.[19]

Strategia a funcționat. Cu cât măsurile poliției deveneau mai inteligente, cu atât metrourile erau mai sigure. Între 1990 și 1992, hărțile lui Maple și tacticile lui Bratton au redus fărădelegile cu 27%, iar jafurile cu o treime.[20]

Când Bratton a devenit comisar al Departamentului de Poliție al orașului New York, a decis să ia cu el „hărțile despre viitor" ale lui Maple. Acolo, acestea au fost dezvoltate și îmbunătățite, devenind CompStat, un instrument de urmărire a informațiilor care este utilizat în prezent de numeroase inspectorate de poliție, atât din Statele Unite, cât și din străinătate. La baza sa se află tot principiul simplu al lui Jack Maple: consemnează unde au avut loc infracțiunile pentru a descoperi cele mai periculoase zone din oraș.

Aceste puncte fierbinți tind să fie, relativ, concentrate. În Boston, de exemplu, un studiu desfășurat pe o perioadă de 28 de ani a descoperit că 66% dintre toate jafurile stradale aveau loc pe doar 8% din străzi.[21] Într-un alt studiu, s-au marcat pe hartă 300 000 de locuri semnalate de apelurile de urgență la poliția din Minneapolis: jumătate dintre ele proveneau de pe o suprafață de numai 3% din întreaga întindere a orașului.[22]

Dar aceste puncte fierbinți nu rămân neschimbate de-a lungul timpului. Ele se deplasează în mod constant, variind și transformându-se subtil, asemenea picăturilor de ulei în apă, chiar și de la o zi la alta. Iar când Bratton s-a mutat la Los Angeles, în 2002, a început să se întrebe dacă mai existau și alte tipare, care ar fi putut să te anunțe *când* urma să aibă

loc o infracțiune, nu numai *unde* avea să fie comisă aceasta. Exista vreo modalitate de a interpreta altfel infracțiunile care fuseseră deja comise? În loc să reacționeze doar după săvârșirea fărădelegilor – motivul frustrării lui Maple – sau să le combată, oare nu exista vreun mod de a le prevedea?

SEMNALUL ȘI INTENSIFICAREA

Când se pune problema de a prezice infracțiunile, cel mai bine este să se înceapă de la analizarea furturilor prin efracție. Din moment ce acestea au loc la o anumită adresă, se știe cu precizie unde s-au petrecut – spre deosebire de furtul din buzunare, spre exemplu (pentru că, în fond, este posibil ca victima să nu observe lipsa telefonului mobil decât atunci când ajunge acasă). Cei mai mulți oameni vor face plângere la poliție dacă au fost tâlhăriți, așa că dispunem de seturi de date foarte detaliate și de bună calitate, care sunt mult mai greu de cules în cazul infracțiunilor legate de traficul de droguri, de exemplu. În plus, oamenii știu adesea destul de bine când le-au fost sparte locuințele (poate atunci când erau la serviciu sau când și-au petrecut seara în oraș), informații de care nimeni nu va dispune în cazul unor infracțiuni precum faptele de vandalism.

Spărgătorii de locuințe au, de asemenea, ceva în comun cu criminalii și violatorii în serie pe care i-a studiat Rossmo: ei tind „să opereze" doar în zone care le sunt familiare. Acum știm că este mai probabil ca o locuință să fie spartă dacă se află pe o stradă des frecventată de un hoț – de pildă, în drumul lui spre locul de muncă sau spre școală.[23] Mai știm* că există un punct sensibil referitor la nivelul de aglomerație

* De această dată, atunci când mă exprim la plural, mă refer la faptul că studiul în cauză e o lucrare pe care am realizat-o împreună cu minunatul meu doctorand, Michael Frith (n. a.).

pe care îl preferă spărgătorii: ei tind să evite străzile cu foarte mult trafic și să le prefere pe acelea pe care se plimbă liniștiți o mulțime de oameni care nu locuiesc în zonă, pentru că aceștia le servesc drept acoperire (atâta vreme cât nu există și mulți localnici băgăcioși care-și pierd vremea pe acolo, făcând pe paznicii improvizați).[24]

Dar acesta este numai primul din cele două elemente care fac o locuință atrăgătoare pentru un spărgător. Este adevărat, există și factori care nu se schimbă în timp, cum ar fi adresa la care locuiți sau cât de aglomerată este strada dumneavoastră, care vor „semnala", vor atrage în mod constant atenția asupra proprietății dumneavoastră, transformând-o într-o țintă. Dar, înainte de a vă grăbi să vă vindeți casa și să vă mutați pe o alee închisă, liniștită, și intens patrulată de poliția comunitară, ar trebui să mai știți că în lumea infracționalității punctele fierbinți nu rămân imobile. Cel de-al doilea motiv pentru care locuința dumneavoastră este atrăgătoare rămâne, fără îndoială, și cel mai important. Acest factor depinde de ceea ce se întâmplă chiar în acest moment în imediata ei vecinătate și este cunoscut drept „intensificare".

Dacă ți-a fost vreodată spartă locuința de două ori într-un interval scurt de timp înseamnă că ești deja bine familiarizat cu efectul de intensificare. Așa cum îți vor spune și polițiștii după ce ai devenit victima unei spargeri pentru prima oară, infractorii au tendința să vizeze în mod repetat același loc – ceea ce înseamnă că, oriunde ai locui, în zilele următoare jafului, există riscul de a fi prădat din nou. De fapt, șansele de a fi vizat din nou cresc de douăsprezece ori în această perioadă.[25]

Sunt câteva motive clare pentru care spărgătorii de locuințe ar putea decide să revină în casa voastră. Poate au ajuns să-i cunoască planul, să afle unde vă păstrați bunurile de valoare (obiecte precum televizoarele și calculatoarele sunt înlocuite destul de repede), să se familiarizeze cu încuietorile ușilor sau căile de scăpare din zonă, ori poate că au observat un obiect

valoros pe care nu au putut să-l ia de prima dată. Oricare ar fi motivul, acest efect de intensificare nu este valabil numai pentru dumneavoastră. Cercetătorii au descoperit că şansele ca vecinii să fie jefuiţi imediat după o spargere vor creşte de asemenea, la fel ca şi cele ale vecinilor vecinilor voştri, ale vecinilor vecinilor vecinilor voştri şi aşa mai departe, pe toată strada.

Vă puteţi imagina cum aceste intensificări îşi fac apariţia şi se aprind în tot oraşul, ca un foc de artificii. Pe măsură ce te îndepărtezi de scânteia iniţială, impulsul devine din ce în ce mai slab, iar, cu timpul, se atenuează treptat până când, după două luni – dacă nu cumva o nouă spargere nu relansează mecanismul pe aceeaşi stradă –, va fi dispărut în totalitate.[26]

Semnalul şi intensificarea întâlnite în fenomenul infracţional au, de fapt, un corespondent într-un fenomen natural: cutremurele de pământ. Este adevărat, nu poţi preciza exact unde şi când va lovi primul cutremur (deşi ştii că unele regiuni sunt mai predispuse decât altele). Dar, în momentul în care se manifestă primele trepidaţii ale solului, poţi să prezici cu destul de multă siguranţă locul şi la frecvenţa replicilor care e de aşteptat să urmeze, riscul fiind cel mai mare acolo unde s-a petrecut cutremurul iniţial, scăzând pe măsură ce te îndepărtezi de el şi estompându-se odată cu trecerea timpului.

Sub conducerea lui Bratton s-a făcut pentru prima oară legătura între structura unui cutremur şi spargerea locuinţelor. Dornic să găsească o cale de a prevedea comiterea infracţiunilor, Departamentul de Poliţie din Los Angeles a încheiat un parteneriat cu un grup de matematicieni de la UCLA* şi i-a lăsat pe aceştia să-şi bage nasul prin toate informaţiile pe care poliţiştii reuşiseră să pună mâna: 13 milioane de infracţiuni comise într-o perioadă de peste 80 de ani. Deşi criminaliştii erau deja de câţiva ani la curent cu fenomenul de semnalare şi intensificare, în încercarea de a

* University of California, Los Angeles (engl., Universitatea din California, statul Los Angeles) (n. tr.).

descoperi anumite pattern-uri în acel noian de informații, cercetătorii UCLA au fost primii care și-au dat seama că ecuațiile matematice care preziceau atât de bine riscul seismic și replicile care apăreau în urma unui cutremur puteau fi, la fel de bine, folosite pentru a prevesti infracțiunile și „post-infracțiunile". Și, în plus, că ele nu dădeau roade numai pentru spargerile din locuințe: ei descoperiseră, practic, o modalitate de a prevedea orice, de la furtul de mașini și până la violențe și fapte de vandalism.

Implicațiile erau cu adevărat... cutremurătoare. În loc să afirmi doar că o zonă din oraș care căzuse recent victimă unor infracțiuni era relativ „mai expusă riscului", cu ajutorul acestor ecuații puteai să cuantifici cu exactitate care era riscul respectiv, ajungând chiar până la nivelul unei singure străzi. Iar știind, probabilistic vorbind, că o anumită zonă din oraș urma să fie ținta spărgătorilor de case într-o anumită noapte, matematicienilor le-a fost ușor să scrie un algoritm care să avertizeze poliția unde să-și concentreze atenția.

Și așa a luat naștere PredPol (sau PREDictive POLicing*).

INFORMATORUL

Se prea poate să fi auzit deja de PredPol. De la lansare, în 2011, a făcut subiectul a mii de articole de ziar, de obicei, sub un titlu care făcea referire la filmul *Minority Report* cu Tom Cruise. A devenit un fel de Kim Kardashian al algoritmilor: foarte celebru, extrem de criticat în mass-media, deși nimeni nu pare să înțeleagă cu adevărat ce face.

Așa că, înainte să vă umpleți mintea cu imaginea unor clarvăzători care stau într-un bazin cu apă și își strigă cu voce

* „Măsuri polițienești preventive" (sau „criminalistică algoritmică") (n. tr.).

tare premoniţiile, daţi-mi voie să vă înşel puţin aşteptările; PredPol nu urmăreşte oamenii înainte ca aceştia să comită o infracţiune; nu poate, în niciun caz, să vizeze un anumit individ, ci doar zone geografice. Ştiu că am utilizat în mod repetat cuvântul „predicţie", însă algoritmul nu poate, de fapt, să prezică exact ce se va întâmpla în viitor. Nu este un glob de cristal. Nu poate prevedea decât riscul apariţiei unor evenimente viitoare, nu şi evenimentele în sine, iar aceasta este o diferenţă subtilă, dar importantă.

Gândiţi-vă la algoritm ca la un fel de agent de pariuri. Dacă un grup numeros de ofiţeri de poliţie se înghesuie în jurul hărţii oraşului, făcând pariuri în privinţa locului în care va fi săvârşită o infracţiune în acea noapte, atunci PredPol calculează şansele ca acest lucru să se întâmple. Acţionează ca un fel de informator, scoţând în evidenţă străzile şi zonele care sunt „favoritele" din noaptea aceea, sub forma unor pătrate mici şi roşii pe o hartă.

Întrebarea-cheie este dacă merită să ţii cont de „favoritul" pe care ţi-l dezvăluie „informatorul". Pentru a verifica dibăcia algoritmului[27], acesta a fost pus să se confrunte cu cei mai buni analişti umani, specialişti în domeniul infracţionalităţii, în cadrul a două experimente diferite: unul realizat în Kent, în partea de sud a Angliei, iar celălalt, în regiunea din sud-vestul oraşului Los Angeles. Testul a constat într-o confruntare cinstită, faţă în faţă. Tot ce aveau de făcut algoritmul şi specialiştii în infracţionalitate era să plaseze pe o hartă 20 de pătrate, fiecare reprezentând o suprafaţă de 150 m^2, pentru a indica unde credeau ei că avea să aibă loc un număr mai mare de infracţiuni în următoarele 12 ore.

Înainte de a vedea care au fost rezultatele, este important să subliniem cât de dificil este acest lucru. Dacă mie sau dumneavoastră ni s-ar da aceeaşi însărcinare, presupunând că nu avem o cunoaştere aprofundată a zonelor cu o rată mare de infracţiuni din Kent sau California, probabil că nu am face decât să aşezăm la întâmplare pătratele pe hartă. Acestea ar

acoperi o porțiune infimă din ea – numai o miime din suprafața totală a comitatului Kent[28] –, iar la fiecare 12 ore ar trebui să eliminăm toate prezumțiile anterioare și să o luăm de la capăt. Împrăștiind pătratele la întâmplare, ne-am putea aștepta „să prezicem" cu succes mai puțin de 1% din infracțiuni.[29]

Specialiștii s-au descurcat mult mai bine. În Los Angeles, analistul a reușit să prezică în mod corect unde vor avea loc 2,1% din infracțiuni[30], iar specialiștii britanici au obținut un procent și mai bun, cu o medie de 5,4%[31], un scor absolut impresionant, ținând cont de faptul că harta lor era cam de zece ori mai mare decât aceea a orașului Los Angeles.

Însă algoritmul i-a eclipsat pe toți. În Los Angeles a prezis corect mai mult decât dublul infracțiunilor pe care reușiseră să le prezică specialiștii umani, iar la o etapă a testului din Marea Britanie, aproape una din cinci infracțiuni s-a produs în pătratele roșii pe care le prevăzuseră ecuațiile matematice.[32] PredPol nu este un glob de cristal pentru ghicit, dar niciuna dintre metodele folosite în trecut nu a reușit să prevadă infracționalitatea viitoare cu un succes atât de mare.

UTILIZAREA PREDICȚIILOR

Numai că apare o problemă. Deși algoritmul este destul de bun la *prezicerea* teritoriului în care vor avea loc infracțiuni în următoarele 12 ore, poliția are un obiectiv puțin diferit: *reducerea* infracționalității în următoarele 12 ore. Odată ce algoritmul a furnizat predicțiile, nu este foarte clar ce ar trebui făcut în continuare.

Desigur, există câteva opțiuni. În cazul spargerilor de locuințe, s-ar putea instala camere CCTV sau s-ar putea aduce polițiști sub acoperire care să-i prindă pe răufăcători în flagrant delict. Sau ar fi mai bine pentru toată lumea dacă toate eforturile s-ar îndrepta spre prevenirea infracțiunilor înainte ca acestea să se întâmple. În definitiv, ce ați prefera?

Să fiți victima unei infracțiuni al cărei făptaș va fi prins sau să nu fiți deloc victima unei infracțiuni?

De asemenea, ar putea fi avertizați locuitorii din zonă că bunurile le sunt în pericol, poate chiar să li se ofere sprijin de către poliție în îmbunătățirea încuietorilor sau s-ar putea instala sisteme de alarmă sau temporizatoare la întrerupătoarele luminilor pentru a păcăli indivizii îndoielnici care ar trece pe stradă, făcându-i să creadă că în locuințe este mereu cineva. Este ceea ce s-a întreprins în cadrul unui experiment din 2012 în Manchester[33], unde s-a reușit reducerea numărului de furturi prin efracție cu peste o treime. A existat un mic dezavantaj, însă: cercetătorii au calculat că această tactică, numită și „asigurarea siguranței țintei", a costat aproximativ 3 925 lire pentru fiecare spargere de locuință preîntâmpinată.[34] Și nu v-aș sfătui să încercați să convingeți Departamentul de Poliție din Los Angeles să pună în aplicare ideea aceasta, căci ofițerii săi se confruntă cu peste 15 000 de spargeri de locuințe pe an![35]

O altă opțiune, care se abate cât mai puțin posibil de la metodele polițienești tradiționale, este tactica numită *cops on the dots**. Iată ce mi-a declarat Steyve Colgan, un ofițer de poliție englez ieșit la pensie care a lucrat la Metropolitan Police**:

> În zilele de demult – patrulele erau repartizate doar pe zone geografice. Luai harta, o fragmentai pe zone și pe acestea le împărțeai pe sectoare. „Tu mergi în sectorul ăla, tu mergi în sectorul ăla!" Simplu ca bună ziua.

Problema era, așa cum a reieșit dintr-un studiu realizat în Marea Britanie, că un ofițer de poliție care patrula pe jos prin sectorul care îi fusese alocat la întâmplare se putea aștepta să

---

* „Polițiști [picați] la țanc" (într-o traducere aproximativă) (n. tr.).

** „Poliția Metropolitană" (engl.), denumire sub care este cunoscută poliția londoneză (oficial: Metropolitan Police Force sau Met), care acționează în cele 32 de burguri londoneze, în afara City-ului, unde operează City of London Police (n. red.).

ajungă la o sută de metri de locul unde avea loc o spargere de locuință numai o singură dată în 80 de ani.[36]

Folosind metoda *cops on the dots*, în schimb, patrulele erau trimise în punctele fierbinți pe care le semnala algoritmul. (Ceea ce înseamnă că metoda ar fi trebuit să se numească, atunci, *cops on the hotspots**, numai că, probabil, denumirea nu ar fi fost la fel de captivantă.) Ideea era, desigur, că, dacă prezența polițiștilor este mai vizibilă, la locul potrivit și la timpul potrivit, e mult mai probabil să oprească desfășurarea unei infracțiuni sau cel puțin să le permită ofițerilor să reacționeze rapid în perioada imediat următoare.

Este exact ce s-a întâmplat în comitatul Kent. În cea de-a doua etapă a studiului, când tocmai începea schimbul de noapte, sergentul a scos la imprimantă hărțile cu pătrate roșii care marcau zonele cu risc ridicat din acea seară. Când patrula de poliție avea o perioadă mai liniștită, se ducea la cel mai apropiat pătrat roșu, cobora din mașină și străbătea zona pe jos.

Într-o seară, într-o zonă în care, în mod normal, nu s-ar fi dus niciodată, polițiștii au găsit pe stradă o femeie din Europa de Est și copilul ei. A reieșit că aceasta era implicată într-o relație abuzivă și că, în urmă cu doar câteva minute, copilul fusese agresat sexual. Sergentul care era de gardă în acea noapte a confirmat că „i-am găsit pe oamenii aceștia pentru că se aflau într-un pătrat PredPol".[37] Suspectul a fost arestat în apropiere de locul acela, mai târziu, în aceeași noapte.

Mama și copilul ei nu au fost singurele persoane ajutate de algoritm pe parcursul experimentului *cops on the dots*: infracționalitatea din Kent a scăzut în medie cu 4%. Studii similare întreprinse în Statele Unite (efectuate chiar de PredPol) au anunțat scăderi și mai importante ale infracționalității. În zona Foothill din Los Angeles infracționalitatea a scăzut cu

* „Polițiști în punctele fierbinți" (engl.) (n. tr.).

13% în primele patru luni de utilizare a algoritmului, în pofida unei creşteri de 0,4% în restul oraşului, unde poliţia întrebuinţa metodele tradiţionale. Iar Alhambra, un oraş din California situat nu foarte departe de Los Angeles, a anunţat o uimitoare scădere de 32% a spargerilor de locuinţe şi de 20% a furturilor de maşini după introducerea algoritmului, în ianuarie 2013.[38]

Aceste cifre sunt impresionante, dar, în realitate, este greu de afirmat cu certitudine dacă meritul este al PredPol. Toby Davies, matematician şi criminolog de la University College London, mi-a spus:

> Este posibil ca până şi simplul fapt că ofiţerii de poliţie sunt încurajaţi să se ducă în diverse locuri, să coboare din maşină şi să patruleze pe jos, indiferent unde, să fi dus oricum la reducere a criminalităţii.

Iar aici mai apare o problemă. În situaţia în care cu cât cauţi mai temeinic infracţiunile, cu atât ai mai multe şanse să le găseşti, atunci faptul de a trimite poliţia pe străzi ar putea să schimbe evidenţa infracţiunilor, căci tot Davies mi-a declarat:

> Când poliţiştii sunt undeva, descoperă mai multe infracţiuni comise acolo decât ar fi făcut-o altminteri. Chiar şi atunci când se petrec, în mod egal, infracţiuni în două locuri diferite, ofiţerii vor descoperi ceva mai multe acolo unde au fost prezenţi, decât vor descoperi în viitor în acela în care nu au fost.

Asta înseamnă că probabil există un dezavantaj uriaş aş metodei *cops on the dots*: trimiţând poliţiştii într-o anumită zonă pentru a combate infracţionalitatea pe baza predicţiilor algoritmului, se riscă intrarea într-un cerc vicios. Pentru că, dacă, spre exemplu, un cartier mai modest are un nivel mai mare de criminalitate într-o primă etapă, algoritmul poate prezice că în viitor acolo se vor produce încă şi mai multe infracţiuni; prin urmare, în acel cartier vor fi trimişi ofiţerii de poliţie, iar aceştia vor descoperi mai multe infracţiuni. Implicit, algoritmul va prezice comiterea şi mai multor infracţiuni, vor fi trimişi acolo încă şi mai mulţi poliţişti şi aşa mai departe.

Acest cerc vicios care se tot repetă s-ar putea să devină o problemă pentru stoparea infracțiunilor corelate cu zonele mai sărace, cum ar fi cerșitul, vagabondajul și consumul de droguri ușoare.

În Marea Britanie, unde unele categorii sociale se plâng în mod constant de lipsa prezenței polițiștilor pe străzi, concentrarea atenției polițiștilor asupra anumitor zone s-ar putea să nu pară inechitabilă imediat. Însă nu toată lumea este în bune relații cu poliția, căci, după cum afirma Davies:

> Este firesc ca persoanele care observă că un ofițer de poliție se plimbă prin fața locuinței lor în fiecare zi să se simtă persecutate din acest motiv, chiar dacă nimeni nu comite nicio infracțiune și chiar dacă polițistul respectiv nu face nimic altceva decât să parcurgă strada de la un capăt la altul. Este aproape ca și când ai avea dreptul să nu te afli constant sub presiune, permanent supravegheat de poliție.

Iar eu, una, tind să-i dau dreptate.

Un algoritm bine pus la punct *ar trebui* să fie astfel dezvoltat, încât să poată ține cont de tacticile utilizate de poliție. Teoretic, cel puțin, există modalități să te asiguri că algoritmul nu vizează cu precădere anumite cartiere (spre exemplu, se pot trimite, în mod aleatoriu, polițiști în zone cu risc mediu și ridicat de comitere a unor infracțiuni). Din păcate, însă nu există nicio modalitate de a ști cu siguranță dacă PredPol reușește să evite cu totul acest cerc vicios care se tot repetă sau dacă operează echitabil la nivel general, întrucât PredPol este un algoritm brevetat, așadar, codul său sursă nu este accesibil publicului și nimeni nu știe exact cum funcționează.

PredPol nu este singurul software existent pe piață. Un concurent al său este HunchLab, care lucrează prin combinarea diferitelor date statistice despre o zonă: infracțiuni declarate, apeluri de urgență, date de recensământ (precum și parametri mai surprinzători cum ar fi fazele lunii). HunchLab nu se bazează pe nicio teorie. Nu încearcă să stabilească *de ce*

infracțiunile au loc mai mult într-o anumită zonă decât în alta; el nu face decât să prezinte pattern-urile pe care le găsește în toate acele informații. În consecință, poate prezice cu fiabilitate mai multe tipuri de infracțiuni decât PredPol (care are la bază teorii despre modul în care infractorii creează tipare geografice) – dar, din moment ce și HunchLab este protejat prin drepturi de proprietate intelectuală, este, practic, imposibil să ne asigurăm, din afară, că nu discriminează, în mod accidental, anumite segmente de populație.[39]

Un alt algoritm predictiv lipsit de transparență este Strategic Subject List*, folosit de Departamentul de Poliție din Chicago.[40] Acesta are o abordare cu totul diferită de a celorlalți: în loc să se concentreze asupra criteriilor geografice, el încearcă să prevadă ce indivizi vor fi implicați în infracțiuni în care se face uz de armă. Utilizând o mare diversitate de factori, algoritmul creează o „listă fierbinte" de persoane despre care consideră că e posibil să fie implicate în acte de violență armată în viitorul apropiat, fie împușcând pe altcineva, fie fiind împușcate. Teoria pe care se bazează este valabilă: victimele de astăzi sunt, adesea, făptașii de mâine. Iar programul este bine intenționat: ofițerii de poliție vizitează persoanele de pe lista de supraveghere pentru a le oferi acces la diverse programe de intervenție și a le ajuta să-și schimbe viața.

Numai că există motive de îngrijorare că Strategic Subject List s-ar putea să nu se ridice la înălțimea așteptărilor. O investigație recentă, realizată de organizația non-profit RAND Corporation** a ajuns la concluzia că faptul de a figura

* „Lista cu subiecți strategici" (engl.) (n. tr.).

** Acronim pentru Research ANd Development Corporation, think-thank dedicat politicilor globale înființat în 1948 de Douglas Aircraft Company (una dintre companiile aviatice care au înflorit datorită contractelor militare obținute în perioada celui de-al Doilea Război Mondial), cu scopul declarat de a oferi capacități de cercetare și analiză economică armatei SUA. În prezent, RAND Corporation este finanțată de guvernul american și de diverse persoane fizice și juridice (n. red.).

pe această listă nu avea nicio importanță din perspectiva probabilității ca un individ să fie implicat într-un schimb de focuri.[41] Însemna, totuși, că era mai probabil să fie arestat. Poate – conchidea raportul – din cauză că ofițerii tratau pur și simplu lista de supraveghere ca pe o listă de suspecți oricând avea loc un schimb de focuri.

Algoritmii predictivi utilizați de poliție sunt, fără îndoială, promițători și oamenii responsabili de crearea lor fac, neîndoielnic, acest lucru cu toată bună credința și cu cele mai bune intenții. Însă îngrijorările privind subiectivismul și discriminarea sunt legitime. Iar în opinia mea aceste chestiuni sunt prea importante pentru construirea unei societăți echitabile pentru ca noi să acceptăm, pur și simplu, asigurări precum aceea că forțele de ordine le vor utiliza într-un mod nepărtinitor. Este unul dintre multele exemple care ne dau de înțeles cât de stringent avem nevoie de specialiști independenți și de un organism de reglementare care să se asigure că binele pe care îl fac algoritmii nu va fi depășit de răul produs de ei.

Iar posibilele efecte dăunătoare ne depășesc capacitățile de predicție. Așa cum am văzut deja într-o serie de alte exemple, există un pericol real ca algoritmii să adauge o notă de autoritate unui rezultat incorect. Iar aici consecințele pot fi dramatice. Doar pentru că un calculator afirmă un lucru nu înseamnă că acesta este și adevărat.

CINE TE CREZI?

În 2014, Steve Talley dormea în locuința lui din South Denver când a auzit o bătaie în ușă.[42] A deschis, iar în pragul ei se afla un bărbat care își cerea scuze că îi lovise din greșeală mașina. Străinul l-a rugat pe Talley să iasă până afară, să vadă despre ce este vorba. El i-a făcut pe plac. Pe când se ghemuia, să evalueze daunele suferite de portiera din partea șoferului a mașinii sale[43], s-a declanșat o grenadă cu efect luminos. Și-au făcut apariția trei bărbați îmbrăcați cu jachete și căști negre, care l-au doborât: un bărbat i-a imobilizat capul, altul –

brațele, pe când un al treilea a început să-l lovească în mod repetat cu patul pistolului.

Talley s-a ales cu răni multiple. Până la sfârșitul serii, suferise leziuni ale nervilor, i se formaseră cheaguri de sânge și se alesese cu o ruptură de penis.[44] „Nici măcar nu știam că un penis poate fi rupt", a declarat el mai târziu unui ziarist de la *The Intercept*. „La un moment dat strigam, pur și simplu, să vină poliția. Pe urmă, mi-am dat seama că aceia care mă băteau erau chiar polițiști."[45]

Steve Talley a fost arestat pentru două jafuri comise la băncile din apropiere. În timpul celui de-al doilea, fusese agresat un ofițer de poliție, motiv pentru care, credea Talley, era tratat cu atâta brutalitate în timpul arestării. „Le-am spus că erau nebuni", a declarat, amintindu-și că le-ar fi strigat polițiștilor: „Ați greșit persoana!"

Talley nu mințea. Arestarea sa era rezultatul asemănării lui uimitoare cu persoana potrivită – adevăratul jefuitor de bănci.

Un om de serviciu, angajat în clădirea în care lucra Talley, informase, inițial, poliția după ce văzuse fotografiile făptașului la știrile locale, însă un specialist FBI examinase ulterior înregistrările camerei CCTV folosind un software de recunoaștere facială[46] și ajunsese la concluzia că „individul care apare în înregistrare pare să fie Talley".[47]

K1 Imaginea lui Steven Talley

Q1 Imaginea filmată de camera video din holul băncii americane

Talley avea un alibi de nezdruncinat, însă din cauza declarației expertului FBI, avea să fie nevoie de mai bine de un an întreg înainte ca numele să-i fie pe deplin reabilitat. Între timp, a fost ținut într-o celulă de maximă securitate aproape două luni, până s-au strâns dovezi suficiente pentru eliberarea lui. În consecință, a fost pus în imposibilitatea de a mai lucra, iar până când calvarul a luat sfârșit, Talley își pierduse deja slujba, casa și dreptul de a-și vedea copiii. Toate acestea ca o consecință directă a acestei false identificări.

ALGORITMI CARE VĂD DUBLU

Algoritmii de recunoaștere facială devin deja un instrument folosit în mod obișnuit în investigațiile polițienești moderne. Când li se prezintă o fotografie, o înregistrare sau o captură de la un aparat de fotografiat 3D, acești algoritmi vor detecta fața care apare, îi vor măsura caracteristicile și o vor compara cu acelea din baza de date cu chipuri cunoscute pentru a determina identitatea persoanei reprezentate acolo.

La Berlin, algoritmii de recunoaștere facială capabili să identifice suspecții de terorism cunoscuți sunt antrenați cu ajutorul mulțimilor de oameni care trec prin gări.[48] În SUA, numai în statul New York acești algoritmi au dus la peste 4 000 de arestări, începând din 2010, *doar* pentru fraudă și furt de identitate.[49] Iar în Marea Britanie, camere video instalate pe vehicule care arată ca niște mașini care lucrează pentru Google StreetView străbat în prezent orașele, verificând, în mod automat, dacă semănăm sau nu cu indivizi din baza lor de date cu persoane căutate.[50] Aceste dube au obținut primul succes în iunie 2017, după ce au trecut pe lângă un bărbat din South Wales, pe numele căruia fusese emis un mandat de arestare.[51]

Siguranța și securitatea depind adesea de capacitatea noastră de a identifica și de a recunoaște chipurile umane.

Însă a lăsa această sarcină în seama oamenilor poate fi riscant. Să-i luăm, de exemplu, pe ofițerii de la pașapoarte. Într-un studiu recent, menit să reproducă toate condițiile de securitate dintr-un aeroport, acești profesioniști ai recunoașterii chipurilor nu au reușit să observe persoane care aveau acte de identitate greșite în uimitorul procent de 14% din cazuri – și au respins în mod incorect 6% dintre persoanele care aveau un act de identitate absolut corespunzător.[52] Nu știu ce părere aveți voi, însă eu găsesc că aceste procente sunt mai mult decât îngrijorătoare, dacă stau să mă gândesc la numărul de persoane care trec zi de zi prin Aeroportul Heathrow.

Așa cum vom vedea în continuare, algoritmii de recunoaștere facială sunt, cu siguranță, mai buni decât oamenii. Însă, din moment ce sunt folosiți pentru a vâna infractorii, iar consecințele identificării greșite sunt atât de grave, utilizarea lor ridică o întrebare importantă: cât de ușor ar putea fi confundată o persoană cu alta? Câți dintre noi au vreo sosie pitită pe undeva, asemenea lui Steve Talley?

Un studiu din 2015 părea să sugereze că șansele să ai o dublură în viața reală (indiferent dacă este un spărgător de bănci sau orice altceva) sunt incredibil de mici. Teghan Lucas, de la Universitatea din Adelaide, a făcut, cu migală, câte opt măsurători faciale pe fotografiile a 4 000 de persoane și nu a reușit să găsească nicio singură potrivire printre ele, ceea ce a făcut-o să tragă concluzia că șansele ca două persoane să aibă exact aceeași față erau mai mici de una la un trilion.[53] În conformitate cu aceste calcule, Talley fusese îngrozitor de ghinionist. Ținând cont de faptul că geamănul său malefic locuia în apropiere și că întâmplarea făcea ca acesta să fie un infractor, ne-am putea aștepta să treacă încă zeci de mii de ani înainte ca alt suflet amărât să cadă victimă aceleiași experiențe nefericite.

Avem, însă, motive să bănuim că toate numerele acestea nu prea se potrivesc. Deși este destul de dificil să-ți imaginezi că te-ai putea întâlni cu cineva care are exact chipul tău, dovezile

existenței unor persoane neînrudite, care nu se cunosc, dar care arată identic, par să fie mai des întâlnite decât ar sugera cercetarea lui Lucas.

Să-l luăm pe Neil Douglas, de exemplu, care se îmbarca într-un avion cu destinația Irlanda când și-a dat seama că dublura lui era așezată pe scaunul său! Selfie-ul pe care și l-au făcut, cu toți pasagerii din avion râzând la unison în fundal, a devenit repede viral și, în scurt timp, bărbați roșcați și cu barbă din întreaga lume își trimiteau propriile poze, să demonstreze că și ei aveau o înfățișare asemănătoare. „Cred chiar că la un moment dat formam o mică armată", a declarat Neil pentru postul de televiziune BBC.[54]

Chiar și eu am o istorie personală de adăugat la toate celelalte. Când aveam 22 de ani, un prieten mi-a arătat o fotografie pe care o văzuse pe pagina de MySpace a unei formații muzicale din partea locului. Era un colaj de fotografii făcute în timpul unui concert la care nu participasem, în care apăreau câteva persoane care se distrau, iar una dintre acestea părea ciudat de familiară. Doar ca să mă asigur că nu-mi pierdusem, fără să știu, cunoștința într-o noapte și mă rătăcisem la o petrecere la care nu-mi aminteam să fi

participat, i-am trimis un e-mail solistului formației, care a confirmat ceea ce bănuiam și eu: sosia mea iubitoare de synthpop avea o viață socială mai activă decât a mea.

Așadar, iată-ne, pe Talley, Douglas și pe mine, având fiecare cel puțin câte o sosie, dacă nu chiar mai multe. Suntem trei inși dintr-o populație de șapte miliarde și jumătate de oameni și suntem deja cu mult sub estimarea lui Lucas, de un caz la un trilion.

Există un motiv pentru această discrepanță. Totul se reduce la definirea de către cercetători a noțiunii de „identic". Studiul lui Lucas cerea ca măsurătorile a două persoane să coincidă perfect. Deși Neil și sosia lui sunt incredibil de asemănători, dacă o nară sau lobul unei urechi s-ar deosebi fie și numai cu un milimetru, ei nu ar mai fi considerați sosii 100% prin prisma criteriilor stabilite de cercetătoare.

Dar chiar și atunci când compari două imagini ale aceleiași persoane, măsurătorile exacte nu vor reflecta felul în care fiecare dintre noi ne schimbăm în permanență prin înaintarea în vârstă, boală, oboseală, expresiile pe care le adoptăm sau felul în care chipurile ne sunt distorsionate de unghiul camerei video. Dacă se încearcă surprinderea particularităților unei fețe doar prin măsurarea acesteia în milimetri, se va descoperi tot atâta variație în chipul unei singure persoane, ca și între două persoane diferite. Mai simplu spus: măsurătorile nu sunt suficiente pentru a distinge un chip de altul.

Deși s-ar putea să nu fie perfect identici, îmi pot imagina fără dificultate că i-aș putea confunda pe Neil și pe geamănul său necunoscut în acea fotografie. Ca și în cazul lui Talley – bietul Steve nici măcar nu semăna *chiar* atât de mult cu adevăratul infractor și, cu toate acestea, imaginile au fost interpretate greșit de specialiștii FBI, așa că el a fost acuzat de o infracțiune pe care nu o comisese, fiind aruncat într-o celulă de maximă securitate.

Așa cum au demonstrat ofițerii de la pașapoarte, este uimitor de ușor să confunzi chipuri care nu-ți sunt familiare,

chiar și atunci când acestea sunt doar vag asemănătoare, iar asta înseamnă că noi, oamenii, suntem surprinzător de incapabili să recunoaștem, să identificăm, persoanele necunoscute. Este și motivul pentru care o prietenă de-ale mele susținea că abia a reușit să reziste până la sfârșitul frumosului film al lui Cristopher Nolan, *Dunkirk*, pentru că se chinuia să deosebească actorii între ei. De aceea consideră adolescenții util „să împrumute" cartea de identitate a unui prieten cu câțiva ani mai mare, ca să cumpere băuturi alcoolice. Și acesta este motivul pentru care Innocence Project, un ONG din SUA, estimează că identificarea greșită de către martorii oculari joacă un rol semnificativ în peste 70% dintre condamnările pe nedrept.[55]

Cu toate acestea, deși un martor ocular ar putea să-l confunde cu ușurință pe Neil cu tovarășul lui de călătorie, mama sa nu ar avea, cu siguranță, nicio dificultate în a-și identifica fiul în acea fotografie. Când este vorba despre persoane cunoscute, ne pricepem extrem de bine să le recunoaștem fizionomiile, chiar și atunci când este vorba despre sosiile din viața reală: poți confunda cu ușurință doi gemeni identici dacă nu-ți sunt decât simple cunoștințe, însă îi vei putea deosebi fără dificultate odată ce ajungi să îi cunoști mai bine.

Iar aici ajungem într-un punct critic: „nu-s asemănătoare lucrurile similare, ci-s asemănătoarea acelea care-mi par mie" astfel, ca să parafrazez un proverb bine-cunoscut. În lipsa unei definiții foarte clare a asemănării, nu poți să măsori *cât* de diferite sunt două fețe și nu există nicio valoare în funcție de care să judecăm pentru a afirma indubitabil că două chipuri sunt identice. Nu poți să definești ce înseamnă să fii o sosie sau să spui cât de comună este o anumită față și nici (ceea ce este extrem de important) nu poți să stabilești care este probabilitatea ca în două imagini să apară același individ.

Asta înseamnă că recunoașterea facială, ca metodă de identificare, nu este asemenea ADN-ului, care stă semeț pe un

fundament statistic solid. Când în criminalistică se folosește testarea ADN-ului, profilarea se concentrează asupra anumitor bucăți din genom despre care se știe că variază foarte mult de la un individ la altul. Amploarea acestei variații ne dă răspunsul: dacă secvența de ADN dintr-un eșantion de țesut uman găsit la locul unei infracțiuni se potrivește cu secvența din proba de salivă a unui suspect, înseamnă că poți să calculezi probabilitatea ca ambele să provină de la aceeași persoană. Mai înseamnă și că poți să indici cu exactitate și care sunt șansele ca o ființă ghinionistă să aibă, întâmplător, o secvență de ADN identică sub aceleași aspecte.[56] Cu cât folosești mai mulți markeri, cu atât sunt mai reduse șansele unei identificări greșite; astfel, alegând numărul de markeri pe care să-i testeze, orice sistem judiciar din lume are deplina putere de a decide asupra nivelului de dubiu pe care este dispus să-l tolereze.[57]

Chiar dacă fețele noastre par atât de intrinsec legate de cine suntem, fără să cunoaștem variația care se manifestă între indivizi, practica identificării infractorilor după chipurile lor nu are o bază științifică riguroasă. Când se pune problema identificării persoanelor după fotografii – ca să citățm o prezentare făcută de o echipă de criminaliști ai FBI: „Lipsa statisticilor înseamnă că, de fapt, concluziile se bazează, în ultimă instanță, pe opinii personale".[58]

Din nefericire, utilizarea algoritmilor pentru recunoașterea facială nu soluționează această dilemă morală, ceea ce este un motiv foarte bun să manifestăm prudență atunci când îi folosim pentru identificarea infractorilor. Asemănarea și identitatea nu sunt același lucru și nici nu vor fi vreodată, oricât de precis ar deveni algoritmul.

Și mai există un motiv întemeiat pentru a proceda cu atenție atunci când folosim algoritmi de recunoaștere facială. Ei nu sunt chiar atât de buni la recunoașterea chipurilor cum s-ar putea crede.

Algoritmii funcționează folosind una din cele două abordări principale. Primul tip de algoritm construiește un model 3D al unei fețe fie prin combinarea unei serii de imagini 2D, fie prin scanarea cu ajutorul unei camere video speciale, cu infraroșu. Este metoda adoptată de sistemul Face ID, pe care Apple îl folosește la iPhone-urile sale. Algoritmii aceștia au găsit o cale să ocolească problemele puse de diferitele expresii faciale și de înaintarea în vârstă, concentrându-se asupra zonelor feței care au țesut rigid și formațiuni osoase, cum ar fi curbura orbitei ochiului sau septul nazal.

Apple a susținut că șansele ca o persoană oarecare să-ți deblocheze telefonul care utilizează Face ID sunt de una la un milion, însă algoritmul nu funcționează absolut impecabil. El poate fi păcălit de gemeni[59], de frați sau surori[60] și de către copiii care șterpelesc telefoanelor părinților lor. (Curând după lansarea Face ID, pe internet a apărut o înregistrare video cu un băiat de zece ani care a reușit să păcălească sistemul de recunoaștere facială de pe iPhone-ul mamei sale. Acum aceasta își șterge SMS-urile dacă în ele apare ceva ce nu ar dori să fie văzut de către fiul ei.)[61] S-au mai relatat cazuri în care algoritmul a fost păcălit de o mască specială, printată la o imprimantă 3D, cu niște imagini infraroșii lipite în locul ochilor.[62] Toate acestea înseamnă că, deși algoritmul ar putea fi destul de bun ca să-ți blocheze telefonul, probabil că nu este suficient de performant ca să fie utilizat pentru a-ți permite accesul la conturile bancare.

Acești algoritmi 3D nu sunt de prea mare folos nici la scanarea fotografiilor de pașaport sau a înregistrărilor camerelor CCTV.. Pentru asta ai nevoie de un al doilea tip de algoritm, care se limitează la imaginile 2D și folosește o abordare statistică. Acești algoritmi nu au în vedere neapărat reperele pe care dumneavoastră sau eu le-am putea recunoaște ca trăsături distinctive, ci construiesc o descriere

statistică a tiparelor cu zone luminoase și întunecate dintr-o imagine. Asemenea algoritmilor din capitolul „Medicina", construiți să recunoască un câine într-o imagine, cercetătorii și-au dat de curând seama că, în loc să te bazezi pe oameni care să decidă ce tipare vor funcționa mai bine, poți să faci algoritmul să învețe care sunt cele mai bune combinații prin metoda încercării și a erorii, aplicată la o imensă bază de date cu fețe umane. De regulă, acest lucru se realizează utilizând rețele neurale. Cu genul acesta de algoritmi s-au făcut, recent, progrese importante în materie de performanță și acurateță. Această performanță, însă, are și un preț. Nu este întotdeauna clar *cum* decide algoritmul dacă o față este asemenea alteia.

Înseamnă, totodată, și că acești algoritmi de ultimă generație pot fi relativ ușor păcăliți. Din moment ce funcționează prin detectarea unei descrieri statistice a tiparelor de lumină și umbră de pe o față, îi poți păcăli chiar și prin purtarea unor ochelari mai zbanghii pe care e imprimat un model aiurit și care induce în eroare. Și mai bine încă: prin crearea unui model năucitor specific menit să indice chipul altei persoane poți, efectiv, să faci algoritmul să creadă că ești acea persoană – așa cum procedează bărbatul din imaginea următoare, care poartă niște ochelari ce îl fac să „arate ca" actrița Milla Jovovich.[63] Adică poți să folosești ochelarii să te deghizezi? Clark Kent știa el ce știa!

Dar, făcând abstracție de atacurile personalizate cu ochelari caraghioși, abilitățile de recunoaștere ale acestor algoritmi statistici au stârnit numeroase titluri admirative în mass-media, precum cele care au întâmpinat apariția FaceNet al lui Google. Pentru a-i pune la încercare abilitățile de recunoaștere, lui FaceNet i s-a cerut să identifice 5 000 de imagini cu chipurile unor celebrități. Persoane specializate în acest tip de recunoaște încercaseră anterior să îndeplinească aceeași sarcină și se descurcaseră excepțional de bine, obținând un procent de 97,5% de identificări corecte (deloc surprinzător, din moment ce chipurile acestor oameni celebri le erau deja familiare participanților).[64] FaceNet s-a descurcat chiar mai bine, obținând procentul fenomenal de 99,6% de identificări corecte.

La prima vedere, s-ar părea că mașinile acestea dețin niște capacități supraumane de recunoaștere a fețelor. Asta lasă impresia unui rezultat de excepție, fără îndoială, suficient de bun pentru a justifica utilizarea algoritmilor pentru identificarea infractorilor. Numai că aici se ascunde o capcană. Cinci mii de fețe înseamnă, în realitate, o cifră deplorabil de mică pe care să-ți testezi algoritmul. Dacă acesta urmează să fie pus la lucru pentru combaterea infracționalității, va trebui să găsească o față din milioane, nu din mii de alte fețe.

Asta pentru că poliția britanică deține acum o bază de date de 19 milioane de imagini cu fețele noastre, creată din toate acele fotografii ale indivizilor arestați în baza suspiciunii de a fi comis o infracțiune. Totodată, FBI are o bază de date de 411 milioane de imagini în care se afirmă că ar apărea jumătate din toți americanii adulți.[65] Iar în China, unde baza de date a cărților de identitate emise oferă un acces facil la miliarde de fețe, autoritățile au investit deja masiv în recunoașterea facială. Există camere de supraveghere instalate pe străzi, în metrouri și aeroporturi, care se presupune că vor detecta orice, de la infractori căutați de poliție până la pietoni care

traversează strada prin locuri nepermise, prin diverse orașe din țară.[66] (Există chiar o sugestie ca și abaterile minore ale cetățenilor din lumea reală, cum ar fi aruncarea gunoiului pe jos, să fie incluse în sistemul de punctaj Sesame Credit, ceea ar atrage asupra acestora toate pedepsele pe care le-am dezvăluit la capitolul „Datele personale informatizate".)

Iată unde este problema: șansele de identificare greșită se înmulțesc spectaculos odată cu numărul fețelor adăugate. Cu cât sunt mai numeroase fețele printre care caută algoritmul, cu atât cresc șansele ca acesta să găsească două fețe care să arate la fel. Prin urmare, în momentul în care încerci să folosești aceiași algoritmi în cataloage care conțin tot mai multe fețe, precizia lor scade.

Ar fi ca și când mi-ați cere să potrivesc ce carte de identitate îi corespunde fiecăreia dintre zece persoane necunoscute, iar atunci când am obținut punctajul maxim să susțineți că am reușit să identific corect fețele în 100% din cazuri, după care să mă lăsați să mă plimb la întâmplare prin centrul New York-ului, să identific infractorii recunoscuți. În mod inevitabil, acuratețea mea va scădea.

Același lucru se întâmplă și cu algoritmii. În anul 2015, Universitatea din Washington a organizat așa-numita competiție MegaFace, în care oameni din toată lumea au fost invitați să își testeze algoritmii de recunoaștere facială pe o bază de date de un milion de fețe.[67] Deși aceasta bază de date se apropie, ca dimensiune, de cataloagele deținute de unele autorități guvernamentale, ea tot este considerabil mai mică decât aceștia. Chiar și așa, algoritmii nu s-au descurcat prea bine în această competiție.

FaceNet al companiei Google – care funcționase aproape perfect în cazul celebrităților, dintr-odată, a obținut doar 75%* ca rată de identificare.[68] Alți algoritmi au ajuns la o rată de

* Această cifră se referă la cât de multe potriviri reale de chipuri a omis algoritmul atunci când a fost setat să evite identificările greșite. În capitolul „Justiția" sunt descrise mai multe tipuri de erori pe care

succes de-a dreptul jalnică, de numai 10%. În momentul în care scriu această carte, cea mai bună ofertă din lume este aceea a chinezilor, un algoritm numit Tencent YouTu Lab, care reușește să obțină o rată de recunoaștere de 83,29%.[69]

Altfel spus, judecând după aceste procente, dacă te-ai afla în căutarea unui anumit infractor într-o listă digitală de milioane de fețe, în cel mai bun caz, în unul din șase cazuri nu vei găsi persoana potrivită.

Ar trebui să mai adaug că acest domeniu progresează foarte repede. Ratele de identificare corectă sunt în creștere ușoară și nimeni nu poate să spună, cu certitudine, ce se va întâmpla în lunile sau în anii care vor urma. Ceea ce pot să afirm este că diferențele de luminozitate, postură, calitate a imaginii și aspect general fac ca recunoașterea cu acuratețe și fiabilitate a fețelor să fie o problemă într-adevăr foarte complicată. Suntem oarecum departe de a obține o precizie perfectă în identificare atunci când căutările se fac într-o bază de date de 411 milioane de fețe sau de a reuși să găsim acea potrivire de unu la un trilion în cazul sosiilor.

STABILIREA UNUI ECHILIBRU

Acestea sunt realități, dar nu neapărat și motive de renunțare. Există algoritmi destul de buni pentru a fi folosiți în anumite situații. În Ontario, de exemplu, persoanele care suferă de dependență de jocurile de noroc se pot înscrie voluntar pe o listă în virtutea căreia li se interzice accesul în cazinouri. Dacă hotărârea lor de a rămâne pe drumul cel drept nu mai e fermă, chipul le va fi descoperit de algoritmii de recunoaștere ai cazinourilor, care îi vor alerta pe angajați să le ceară în mod politicos pariorilor să părăsească incinta.[70] Sistemul este, cu siguranță, nedrept față de toți cei care sunt împiedicați din

le poate face un algoritm, iar diferitele modalități de măsurare a acurateței sunt prezentate în capitolul „Medicina" (n. a.).

greşeală să petreacă o noapte la jocul de ruletă, însă cred că acest preţ merită plătit dacă asta înseamnă să ajuţi o persoană dependentă de jocuri de noroc, aflată în plin proces de renunţare la viciu, să reziste tentaţiei de a-şi relua vechile apucături.

Acelaşi lucru este valabil şi în cazul vânzării cu amănuntul. Personalul de pază din magazine obişnuia să tapeteze pereţii birourilor cu fotografii Polaroid ale hoţilor din magazine; în prezent, algoritmii pot să-ţi compare faţa cu acelea din baza de date cu hoţii cunoscuţi chiar din clipa în care ai trecut pragul magazinului. Dacă faţa ta se potriveşte cu aceea a unui infractor bine-cunoscut, este trimisă o alertă pe smartphone paznicilor care te pot găsi apoi prin raioane.

Magazinele au o motivaţie puternică să îşi dorească utilizarea acestui gen de tehnologie. Aproximativ 3,6 milioane de infracţiuni sunt comise anual în magazine numai în Marea Britanie, costându-i pe comercianţi uluitoarea sumă de 660 de milioane de lire sterline.[71] Când te gândeşti că în anul 2016 au fost omorâţi violent 91 de oameni suspectaţi că ar fi furat din magazine din SUA[72], acesta este un argument în favoarea faptului că o metodă de a-i împiedica pe infractorii insistenţi să intre într-un magazin înainte ca situaţia să escaladeze ar fi benefică pentru toată lumea.

Numai că această soluţie high-tech la furtul din magazine are şi dezavantaje: în primul rând, confidenţialitatea (Face First, unul dintre furnizorii importanţi al acestui tip de software de securitate, susţine că nu stochează imaginile clienţilor obişnuiţi, însă magazinele folosesc, cu siguranţă, recunoaşterea facială pentru ne urmări şi a ne analiza comportamentul la shopping). Iar apoi se mai pune problema: cine ajunge să figureze pe lista neagră digitală. De unde ştim că toţi cei de pe listă se află acolo din motive întemeiate? Cum rămâne cu prezumţia de nevinovăţie? Ce se întâmplă cu cei care ajung pe listă în mod accidental: cum vor reuşi să iasă de acolo? La toate acestea se adaugă şi faptul că apare

posibilitatea identificării greşite de către un algoritm, care nu poate fi niciodată absolut precis.

Întrebarea care se pune este dacă avantajele depăşesc dezavantajele. Răspunsul nu este simplu. Nici măcar comercianţii nu sunt, în unanimitate, de aceeaşi părere. Unii adoptă cu entuziasm această tehnologie, pe când alţii o resping – inclusiv Walmart, care a pus capăt testării algoritmului FaceFirst în magazinele sale după ce nu a dat randamentul pe care-l spera compania în urma investiţiei în acesta.[73]

În cazul infracţionalităţii, însă, echilibrul dintre rău şi bine pare mult mai clar delimitat. Este adevărat, aceşti algoritmi nu sunt singurii utilizaţi, căci fundamentările lor statistice sunt cam precare. Amprentarea nu are nicio rată de eroare cunoscută[74], nici analiza urmelor de muşcătură, nici tiparul stropilor de sânge[75] şi nici balistica[76]. De fapt, dacă ar fi să dăm crezare unei lucrări din 2009 a Academiei Naţionale de Ştiinţe a SUA, niciuna dintre tehnicile de investigaţie criminalistică, exceptând testarea ADN, nu poate „să demonstreze existenţa unei legături între o probă şi un anumit individ sau o anumită sursă".[77] Totuşi, nimeni nu poate să nege că toate s-au dovedit nişte instrumente extraordinar de utile în munca de poliţie – atât timp cât nimeni nu se bazează foarte mult pe probele pe care le generează. Însă ratele de acurateţe chiar şi ale celor mai sofisticaţi algoritmi de recunoaştere facială lasă mult de dorit. Se susţine că, dacă există chiar şi cel mai mic risc să apară mai multe cazuri precum acela al lui Steve Talley, atunci o tehnologie imperfectă nu ar trebui utilizată pentru a contribui la răpirea libertăţii unei persoane. Singura problemă este că poveşti ca aceea a lui Talley nu prea ajung să descrie întregul tablou. Pentru că, deşi utilizarea recunoaşterii faciale pentru prinderea infractorilor are dezavantaje uriaşe, ea are şi avantaje imense.

UN COMPROMIS DELICAT

În mai 2015, un bărbat fugea pe străzile din Manhattan atacând la întâmplare trecătorii cu un ciocan mare şi negru.

La început, s-a îndreptat în fugă către un grup de oameni de lângă Empire State Building și l-a lovit cu putere în ceafă pe un tânăr de 20 de ani. Șase ore mai târziu, s-a îndreptat spre sud, spre Union Square, și cu același ciocan a atacat o femeie care stătea liniștită pe o bancă din parc, lovind-o în tâmplă. Doar câteva minute mai târziu, și-a făcut din nou apariția, vizând de data asta o femeie de 33 de ani care mergea pe o stradă care mărginea parcul.[78] Folosind înregistrările camerelor de supraveghere care surprinseseră atacurile, un algoritm de recunoaștere facială a reușit să-l identifice pe David Baril, un bărbat care postase pe Instagram, cu câteva luni înaintea atacurilor, o poză cu un ciocan de pe care picura sânge.[79] Acesta s-a declarat vinovat de acuzațiile care rezultau din atacuri și a fost condamnat la 22 de ani de închisoare.

Și cazurile vechi sunt readuse la viață de progresele recunoașterii faciale. În 2014, un algoritm a adus în fața justiției un bărbat de cetățenie americană care trăia de 15 ani ca un fugar, sub un nume fals. Neil Stammer se sustrăsese de la executarea condamnării sale la închisoare în timpul eliberării pe cauțiune pentru acuzații care includeau abuz sexual asupra minorilor și răpire; a fost arestat din nou atunci când poza lui de persoană căutată de FBI a fost comparată din nou cu baza de date a pașapoartelor și a fost găsită o potrivire cu o persoană care locuia în Nepal, a cărei fotografie de pașaport purta un nume diferit.[80]

După vara anului 2017, când au murit opt persoane în urma unui atac terorist pe London Bridge, îmi dau seama cât de util poate fi un sistem care ar folosi un astfel de algoritm. Youssef Zaghba a fost unul dintre cei trei bărbați care au intrat cu mașina în pietoni, înainte de a porni să înjunghie oameni la întâmplare în apropiere, în Borough Market. Figura pe lista persoanelor suspectate de activități teroriste din Italia și ar fi putut fi identificat automat de un algoritm de recunoaștere facială înainte de a intra în Marea Britanie.

Dar cum decizi să faci acel compromis între confidențialitate și protecție, între echitate față de toată lumea și siguranță? Cât de mulți Steve Talley suntem dispuși să acceptăm în schimbul identificării rapide a unora precum David Baril și Youssef Zaghba?

Să aruncăm o privire asupra statisticilor oferite de NYPD. În anul 2015, acest departament de poliție a raportat identificarea cu succes a 1 700 de suspecți, ceea ce a dus la 900 de arestări, dar, totodată, a identificat greșit cinci persoane.[81] Oricât ar fi de îngrijorătoare eroarea făcută cu fiecare dintre cele cinci, rămâne întrebarea: este acesta un procent acceptabil? Este un preț pe care suntem dispuși să-l plătim pentru a reduce rata criminalității?

După cum se pare, algoritmii lipsiți de dezavantaje precum geo-profilarea lui Kim Rossmo, despre care am discutat la începutul capitolului, reprezintă, mai degrabă, excepția decât regula. Atunci când se ajunge la combaterea infracționalității, oriunde ai căuta, vei găsi algoritmi foarte promițători sub un anumit aspect, dar care pot fi foarte îngrijorători sub un altul. PredPol, HunchLab, Strategic Subject List și recunoașterea facială – toate promit să ne rezolve problemele, însă toate creează altele noi pe parcurs.

După părerea mea, necesitatea reglementării urgente a utilizării algoritmilor nu a fost niciodată mai acută decât în cazul infracționalității, unde simpla existență a acestor sisteme informatice ridică întrebări serioase pentru care nu există răspunsuri facile. Va trebui să ne confruntăm cumva cu aceste dileme dificile. Ar fi bine să acceptăm numai algoritmii pe care îi putem înțelege sau înăuntrul cărora putem arunca o privire, știind că a-i lua de sub oblăduirea proprietarilor lor ar putea duce la pierderea eficienței (iar rata infracționalității va crește)? Sau să respingem din start orice sistem matematic care are incluse din construcție unele scăpări sau care s-a dovedit capabil de eroare, știind că în felul acesta vom impune

algoritmilor un standard mai înalt decât îi cerem societății în care trăim? Și cât de tendențios înseamnă „prea tendențios"? Când să acordăm prioritate victimelor unor infracțiuni care puteau fi prevăzute, în defavoarea victimelor pe care le face algoritmul?

Într-o anumită măsură, asta se reduce la a decide, ca societate, ce credem noi că ar însemna succesul. Care este prioritatea noastră? Să menținem infracționalitatea la cel mai scăzut nivel cu putință? Sau să protejăm, mai presus de orice, libertatea celor nevinovați? Cât de mult dintr-una ați fi dispuși să sacrificați de dragul celeilalte?

Gary Marx, profesor de sociologie la MIT, a descris foarte bine această dilemă într-un interviu pe care l-a dat revistei *The Guardian*, în care afirma:

> Uniunea Sovietică avea un nivel remarcabil de scăzut de delicte minore de stradă în perioada în care controlul totalitar și autoritar era cel mai atroce. Dar, Dumnezeule, cu ce preț obținuse asta?[82]

Este posibil să decidem, până la urmă, că ar trebui să le impunem algoritmilor anumite limite până la care să aibă acces, pentru ca anumite lucruri să nu fie analizate și calculate. Această perspectivă s-ar putea aplica și dincolo de lumea infracționalității. Nu neapărat din cauza lipsei de străduință din partea algoritmilor înșiși, ci poate – doar „poate" – pentru că există și lucruri care depășesc competența unei mașini lipsite de simțire.

Capitolul 7

ARTA

Justin căzuse pe gânduri. Pe 4 februarie 2018, în sufrageria locuinței lui din Memphis, statul Tennessee, urmărea Super Bowl, meciul de super-cupă de fotbal american și mânca bombonele M&M. Mai devreme, în aceeași săptămână, sărbătorise împlinirea a 37 de ani, iar acum – ceva ce devenise pentru el o tradiție anuală – cugeta la ce se alesese de viața lui.

Știa că ar trebui să fie recunoscător, chiar știa asta. Avea o viață cât se poate de confortabilă. O slujbă de birou de la nouă la cinci, un acoperiș deasupra capului și o familie iubitoare. Numai că își dorise dintotdeauna ceva mai mult. În copilărie și adolescență, fusese tot timpul convins că era menit să aibă faimă și avere.

Cum ajunsese, atunci, să fie atât de... normal? „Totul mi s-a tras de la trupa aceea de băieți", se gândea el, aceea în care intrase la 14 ani. „Dacă am fi avut succes, acum totul ar fi fost diferit." Dar din cine știe ce motiv, trupa fusese un fiasco. Succesul nu se prea arătase pentru bietul, sărmanul Justin Timberlake.

Abătut, mai deschise o bere și își imagină cum ar fi putut să fie viața sa. Pe ecranul televizorului, reclamele de la Super Bowl luară sfârșit. Începu muzica, anunțând marele spectacol al reprizei a doua. Iar într-un univers paralel – practic identic cu acesta cu excepția unui singur

detaliu – un alt Justin Timberlake de 37 de ani, din Memphis, intră în scenă.

MAI MULTE LUMI

De ce adevăratul Justin Timberlake avea atât de mult succes? Şi de ce gânditorul Justin Timberlake eşuase? Unii (inclusiv eu, cea care am fost la vârsta de 14 ani)* ar putea susţine că succesul lui Justin este binemeritat: înzestrarea nativă, aspectul fizic plăcut, talentul de dansator şi valoarea artistică a muzicii sale au făcut din celebritate o certitudine. Alţii, însă, ar putea să nu fie de acord. Poate vor susţine că Timberlake nu are nimic deosebit, ca, de altfel, oricare alt superstar al muzicii pop adulat de legiuni întregi de fani. Este uşor să găseşti oameni talentaţi care se pricep să cânte şi să danseze – starurile nu sunt decât cei care au avut noroc.

Bineînţeles că nimeni nu poate fi sigur de asta. Nu fără să construiască un număr de lumi identice paralele, aducându-l pe Timberlake în fiecare dintre ele şi urmărind evoluţia tuturor întrupărilor sale ca să vadă dacă are succes de fiecare dată. Din păcate, crearea unui multivers artificial ne depăşeşte pe mulţi dintre noi, însă, dacă ne situăm sub nivelul lui Timberlake şi luăm în considerare muzicienii mai puţin cunoscuţi, se poate studia rolul relativ pe care îl joacă şansa şi talentul în popularitatea unui hit.

Exact asta a fost ideea care a stat la baza unui experiment celebru, realizat de cercetătorii Matthew Salganik, Peter Dodds şi Duncan Watts în 2006, în care au fost create mai multe lumi digitale.[1] Oamenii de ştiinţă şi-au construit propriul player audio on-line, un fel de versiune foarte primitivă a Spotify, şi au filtrat vizitatorii spre opt website-uri paralele de muzică, pe

* Pe de altă parte, eul meu de 14 ani era un mare fan al duoului PJ & Duncan, aşa că nu prea avea habar de multe! (n. a.).

fiecare din ele fiind instalate aceleași 48 de cântece ale unor artiști necunoscuți.

În ceea ce a devenit cunoscut drept *the Music Lab*[2], 14 341 de pasionați de muzică au fost invitați să se autentifice în player, să asculte clipurile de pe fiecare playlist, să evalueze cântecele și să descarce melodia care le plăcea mai mult.

La fel ca în Spotify-ul real, vizitatorii puteau vedea dintr-o privire ce muzică ascultau alte persoane în „lumea" lor. Alături de numele artistului și de titlul cântecului, participanții vedeau de câte ori fusese descărcată piesa respectivă în „lumea" lor. Toate contoarele au început de la zero, iar în timp, pe măsură ce cifrele se schimbau, cântecele cele mai apreciate în fiecare dintre cele opt topuri paralele începeau să se detașeze cu claritate.

Între timp, pentru a obține nivelul „realei" popularități a cântecelor, echipa de cercetători a construit și o „lume de control", în care preferințele vizitatorilor nu puteau fi influențate de acelea ale altor utilizatori. Aici, cântecele aveau să apară pe pagină într-o ordine aleatorie – într-o grilă sau pe o listă – însă statisticile descărcărilor urmau să fie ascunse vederii.

Rezultatele obținute i-au intrigat pe toți. Toate „lumile" au fost de acord că anumite cântece erau de-a dreptul jalnice. Altele erau veritabile hituri: acestea au sfârșit prin a deveni populare în fiecare lume, chiar și în aceea în care vizitatorii nu puteau să vadă numărul descărcărilor. Dar între hiturile incontestabile și chixurile absolute, artiștii se puteau situa cam la orice nivel de succes.

Să luăm, de exemplu, 52Metro, o formație punk din Milwaukee al cărei cântec, „Lockdown", a fost extrem de popular într-una dintre lumi, unde a ajuns, în final, pe primul loc al topului, dar care, într-o altă lume, a dat chix, situându-se abia pe locul al patruzeci și optulea. Exact același cântec, la concurență cu aceeași listă de cântece; doar că în lumea aceasta 52Metro nu a prins deloc.[3] Se părea că, uneori, succesul chiar e o chestiune de noroc.

Deşi drumul spre primul loc nu era bătut în cuie, cercetătorii au descoperit că era mult mai probabil ca vizitatorii să descarce melodiile despre care ştiau că fuseseră apreciate de către alţi ascultători. Dacă un cântec mediocru ajungea, din întâmplare, în topul preferinţelor mai devreme, popularitatea lui putea să crească asemenea unui adevărat bulgăre de zăpadă. Mai multe descărcări generau încă şi mai multe descărcări. Popularitatea percepută devenea o popularitate reală, aşa încât eventualul succes nu era decât o întâmplare care ajunsese să se amplifice în timp.

Aceste rezultate aveau o motivare: fenomenul cunoscut de către psihologi drept „dovadă socială"* sau „influenţare prin intermediul informaţiilor primite din societate". Ori de câte ori nu dispunem de informaţii suficiente pentru a lua noi înşine decizii, obişnuim să imităm comportamentul celor din jur. De aceea în sălile de spectacol sunt introduse, în secret, printre spectatori, câteva persoane care să aplaude şi să ovaţioneze la momentul potrivit. Imediat ce-i auzim pe alţii aplaudând, este foarte probabil să o facem şi noi. Când vine vorba despre alegerea muzicii, nu înseamnă neapărat că preferăm să ascultăm aceleaşi cântece ca şi alte persoane, ci doar că popularitatea lor este o cale rapidă de a ne feri de dezamăgire. Sau, după cum a declarat Salganik însuşi revistei *LiveScience*:

> Oamenii sunt puşi în faţa prea multor opţiuni. Din moment ce nu poţi să le asculţi pe toate, o scurtătură firească este aceea de a asculta ce ascultă şi alte persoane.[4]

Folosim popularitatea ca indicator al calităţii în toate formele de divertisment. Un studiu din 2007, de exemplu, a cercetat impactul pe care îl are asupra percepţiei publicului despre o anumită carte faptul că aceasta a apărut în topul bestsellerurilor din publicaţia *New York Times*. Exploatând

* *Social proof* (engl.), termen inventat de prof. univ. Robert Cialdini, cunoscut în limba română şi sunt denumirea peiorativă „efect de turmă" (n. red.).

particularitățile modului de alcătuire a acestui clasament, Alan Sorensen, autorul studiului, a urmărit cărțile de succes care ar fi trebuit incluse, în funcție de numărul exemplarelor vândute, însă care – din cauza întârzierilor și omisiunilor accidentale – nu mai apăruseră pe listă și le-a comparat cu acelea care reușiseră să figureze. A descoperit un efect dramatic: simplul fapt că unele cărți figurau pe listă dusese la o creștere a vânzării lor, în medie, cu 13–14% și la o creștere a vânzărilor cu 57% în cazul autorilor aflați la debut.

Cu cât folosim mai multe platforme pentru a vedea ce este popular într-un anumit moment (liste de bestselleruri, clasamentele de pe Amazon, punctajul obținut de un film pe Rotten Tomatoes, topurile hiturilor de pe Spotify), cu atât dovada socială va avea un impact mai mare. Efectul este amplificat și mai mult atunci când suntem bombardați cu milioane de opțiuni, însoțite de campaniile lor de marketing, impactul celebrității creatorului, de publicitatea făcută în mass-media sau cu ajutorul aprecierilor critice și când absolut toate ne reclamă atenția.

Toate acestea înseamnă că, uneori, un cântec penibil poate ajunge în topul preferințelor (și nu spun asta pentru c-aș fi cinică). Prin anii 1990 se zvonea că producătorii britanici de muzică – în deplină cunoștință de cauză – încheiaseră un pariu: care dintre ei reușea să aducă în topurile muzicale cel mai prost cântec cu putință. În urma pariului se spune că a fost aleasă o formație de fete numită Vanilla, al cărei cântec de debut, „No way, no way, mah, na, mah, na" se inspira din cunoscutul cântecel al celebrelor păpuși Muppets. Formația aceasta avea o interpretare care numai cu multă bunăvoință putea fi considerată cântat, o ilustrație muzicală care părea să fi fost desenată cu Microsoft Paint și un video promoțional care avea motive întemeiate să fie considerat cel mai prost realizat din toate timpurile.[5] Însă Vanilla avea niște aliați puternici. Datorită câtorva articole publicate în diverse reviste și al unei

apariții în îndrăgita emisiune *Top of the Pops* a postului BBC, cântecul tot a reușit să ajungă pe locul 14 în topurile muzicale.*

Este de la sine înțeles că succesul formației a fost de scurtă durată. Când a scos cel de-al doilea single, popularitatea ei deja dispăruse. Un al treilea nu a mai apărut. Toate acestea par să sugereze că dovada socială nu este singurul factor care influențează gusturile, așa cum a demonstrat un experiment ulterior al echipei de la Music Lab.

Structura celui de-al doilea studiu a fost, în linii mari, aceeași ca a celui dintâi. Însă, de această dată, pentru a testa în ce măsură percepția referitoare la popularitate devenea o profeție care se auto-împlinea, cercetătorii au adăugat o chichiță. De îndată ce topurile din fiecare „lume" s-au stabilizat, ei au oprit experimentul și au întors lista preferințelor muzicale cu susul în jos. Noii vizitatori ai player-ului audio au văzut hiturile la bază piramidei topului, pe când chixurile de pe ultimele locuri, ajunse în vârf, au început să pară *crème de la crème*** din acel top muzical.

Aproape imediat numărul total de descărcări a început să scadă. În momentul în care cântecele de pe primele locuri nu au mai fost atrăgătoare, oamenii și-au pierdut interesul față de cântecele de pe întregul website. Declinurile cele mai abrupte s-au înregistrat la descărcările fiasco-urilor ajunse acum în vârful topurilor. În acest timp, piesele bune, care lâncezeau pe ultimele locuri, au avut o soartă mai nefericită decât atunci când erau în vârf, dar tot s-au descurcat mai bine decât cele care se situaseră, anterior, la sfârșitul listei de preferințe. Dacă oamenii de știință ar fi lăsat experimentul să continue suficient de mult timp, melodiile cele mai bune dintre cele mai bune și-ar fi restabilit popularitatea. În concluzie: piața

* Este mare păcat că formatul acestei cărți nu permite includerea unor fragmente muzicale, pentru că îmi doresc *foarte mult* să ascultați cât de îngrozitor de prost este cântecul acesta. Căutați-l pe Google, bine? (n. a.).

** „Cele mai bune" (fr. în original) (n. tr.).

nu este blocată într-o anumită stare. Atât șansa, cât și calitatea au un rol important de jucat.[6]

Revenind la realitate – acolo unde s-au strâns suficiente informații doar dintr-o singură lume pe care să le putem folosi – există o interpretare simplă a descoperirilor din experimentele efectuate în Music Lab: calitatea contează și ea nu este totuna cu popularitatea. Faptul că melodiile cele mai bune și-au recâștigat popularitatea dovedește că o anumită muzică este, inerent, „mai bună" decât o alta. La o extremă, un cântec senzațional al unui artist nemaipomenit ar trebui (cel puțin teoretic) să fie menit succesului. Însă problema este că reciproca nu e neapărat valabilă. Numai pentru că un lucru are succes nu înseamnă că este și de calitate.

Definirea calității este o cu totul altă poveste, la care vom ajunge imediat. Însă pentru unii calitatea în sine nu este neapărat importantă. Dacă ești o companie de înregistrări, un producător de film sau un grup editorial, întrebarea de un milion de dolari este: poți să observi cu anticipație un succes garantat? Poate fi creat un algoritm care să selecteze hiturile?

VÂNĂTOAREA DE HITURI

Investirea banilor în filme este o afacere riscantă. Puține sunt filmele care aduc bani; cele mai multe abia dacă își scot cheltuielile, iar eșecurile nu pot fi evitate în totalitate.[7] Este o afacere cu miză mare: atunci când costurile realizării unui film ajung la zeci sau sute de milioane de dolari, eșecul în prezicerea cererii pentru acel produs poate fi catastrofal de costisitor.

Este o lecție pe care au învățat-o pe propria piele cei de la Disney, cu filmul *John Carter*, lansat în 2012. Studioul a aruncat pe apa Sâmbetei 350 de milioane de dolari pentru realizarea sa și a decis că acesta ar fi trebuit să se alăture unor *blockbuster*-e precum *Toy Story* și *Finding Nemo*, devenind primul produs din următoarea lor mare franciză. Cum, nu l-ați văzut? Ei bine, nici eu! Filmul nu a reușit să capteze imaginația

publicului și a sfârșit prin a produce o pierdere de 200 de milioane de dolari, ceea ce a dus la demisia președintelui studiourilor Walt Disney.[8]

Cei mari și tari de la Hollywood au acceptat întotdeauna ideea că efectiv nu ai cum să prezici succesul comercial al unui film. Așa ceva ține de instinct. A-ți încerca norocul cu filme care ar putea să dea lovitura la încasări face parte din meserie. În 1978, Jack Valenti, președinte și director general al Motion Picture Association of America, s-a exprimat în felul următor:

> Nimeni nu-ți poate spune cum va fi primit un film pe piață. Nu înainte să ruleze într-o sală de cinematograf cu luminile stinse, iar între ecran și spectatori să zboare scântei.[9]

Cinci ani mai târziu, în 1983, William Goldman, care a scris romanele care stau la baza filmelor *The Princess Bride* (*File de poveste*) și *Butch Cassidy* și *Sundance Kid*, a fost și mai succint: „Nimeni nu are habar de nimic."[10]

Așa cum am văzut pe parcursul acestei cărți, algoritmii moderni sunt, de obicei, capabili să prezică până și ceea ce pare imposibil de prezis. De ce ar face filmele excepție? Poți să calculezi succesul unui film cu ajutorul venitului obținut din încasări și al aprecierilor critice pe care le obține. Poți să calculezi tot felul de factori despre structura și caracteristicile unui film: distribuția, genul, durata, caracteristicile intrigii și așa mai departe. De ce nu s-ar aplica, așadar, aceleași tehnici în încercarea de a găsi capodoperele? Sau pentru a descoperi ce filme sunt menite să aibă un succes uriaș și încasări pe măsură?

Aceasta a fost ambiția câtorva studii științifice recente, care și-au propus să se inspire din extinsele și bogatele baze de date strânse și organizate de site-uri precum Internet Movie Database (IMDb) sau Rotten Tomatoes. Și – lucru deloc surprinzător! – în aceste informații se ascund nenumărate revelații.

Să luăm, spre exemplu, studiul condus de Sameet Sreenivasan în anul 2013.[11] El și-a dat seama că, deoarece

le ceruse utilizatorilor săi să eticheteze filmele cu ajutorul unor cuvinte-cheie care descriau intriga, IMDb crease un catalog de descriptori uimitor de detaliat care ar fi putut să arate cum se modificaseră preferințele publicului în materie de filme de-a lungul timpului. În momentul în care Sameet și-a început studiul, IMDb catalogase peste 2 000 000 de filme care acopereau un interval de timp de peste 100 de ani, fiecare având mai multe etichete care descriau intriga. Unele cuvinte-cheie erau descrieri mai generale ale filmului, precum „crimă-organizată" sau „relația tată-fiu". Altele se bazau pe localizări, precum „Manhattan, New-York City" sau pe anumite aspecte al intrigii, cum ar fi: „amenințat cu arma" sau „legat de scaun".

De sine stătătoare, cuvintele-cheie arătau că interesul nostru față de anumite elemente ale intrigii tinde să se manifeste în valuri de popularitate (gândiți-vă, spre exemplu, la filmele despre al Doilea Război Mondial sau la acelea care abordează subiectul avortului). Va exista un val de lansări în succesiune rapidă pe teme similare, după care va urma o perioadă scurtă de acalmie. Luând în considerare toate aceste etichete, Sreenivasan a reușit să obțină un punctaj pentru caracterul de noutate al fiecărui film în momentul lansării – un număr cu zecimale, cuprins între zero și unu – care putea fi comparat cu succesul încasărilor sale.

Dacă un anumit aspect al intrigii sau o caracteristică principală (cum ar fi prezența scenelor care prezentau nuduri feminine sau aspecte legate de crima organizată) era un aspect familiar al filmelor mai vechi, cuvântul-cheie avea să-i aducă filmului un punctaj scăzut pentru noutate. Însă caracteristicile originale ale intrigii – cum ar fi introducerea artelor marțiale în filmele de acțiune din anii 1970, să spunem – îi puteau aduce filmului un punctaj ridicat pentru noutate atunci când caracteristica respectivă apărea pentru prima oară pe marile ecrane.

După cât se pare, avem o relație complicată cu acest caracter de noutate. În medie, cu cât punctajul pentru noutate

al unui film era mai ridicat, cu atât acesta avea încasări mai mari. Dar numai până la un punct. Dacă depășeai pragul de noutate, înaintea ta se deschidea o prăpastie: veniturile obținute de un film înregistrau o prăbușire bruscă la orice punctaj care depășea valoarea de 0,8. Studiul lui Sreenivasan dezvăluia ceva ce sociologii bănuiau demult: ne îndepărtăm de ceea ce este banal, dar, totodată, detestăm ceea ce este complet necunoscut. Filmele cele mai bune se situează într-o nișă îngustă și confortabilă, între „nou" și „nu foarte nou".

Punctajul pentru caracterul de noutate poate fi un mod util de a ajuta studiourile de film să evite susținerea unor chixuri, dar nu este de prea mare folos dacă dorești să afli ce soartă va avea un anumit film. Pentru acest lucru ar putea fi mai utilă activitatea unei echipe de cercetători europeni care a descoperit existența unei legături între numărul de editări făcute unui film pe o pagină de Wikipedia în luna dinaintea lansării în cinematografe și posibilele încasări.[12] Editările erau, adesea, făcute de persoane fără legătură cu lansarea filmului, de către simpli pasionați de film care contribuiau cu informații la pagina respectivă. Editările mai numeroase însemnau că în jurul unei lansări se stârnise mai multă vâlvă ceea ce ducea la încasări mai mari.

În ansamblu, modelul lor avea o putere predictivă modestă: din 312 filme studiate, au prezis corect încasările pentru 70 dintre acestea, cu o precizie egală sau mai mare de 70%. Dar cu cât un film era mai bine primit de public, cu atât se făceau mai multe editări pe pagina din Wikipedia și cu atât echipa de cercetători avea mai multe informații pentru a-și continua studiul, iar, prin urmare, cu atât deveneau mai exacte predicțiile pe care le făcea aceasta. Încasările obținute de șase filme cu câștiguri mari au fost prezise corect, cu o acuratețe de 99%!

Studiile acestea sunt interesante sub aspect intelectual, însă un model care funcționează numai cu o lună înainte de lansarea unui film nu este de prea mare folos pentru

investitori. Dar ce ar fi să abordăm întrebarea direct: să luăm toți factorii implicați dinainte cunoscuți – genul filmului, nivelul de celebritate al actorilor principali, clasificarea pe categorii de vârstă (AP, 12 etc.) – și să folosim un algoritm de învățare automată pentru a prezice dacă un film va fi considerat excepțional?

Un studiu celebru din 2005 a făcut exact acest lucru, utilizând o rețea neurală în încercarea de a prezice performanța unor filme cu mult înainte de lansarea lor în cinematografe.[13] Pentru a face lucrurile cât mai simple, autorii săi au renunțat la prezicerea exactă a încasărilor, încercând în schimb să clasifice filmele într-una din nouă categorii, de la eșec răsunător la film cu încasări record. Din păcate, chiar și cu această etapă simplificatoare, rezultatele obținute au lăsat mult de dorit. Rețeaua neurală a depășit performanțele tuturor modelelor statistice folosite anterior, dar tot a reușit să clasifice corect performanța unui film, în medie, în numai 36,9% din cazuri. În categoria de top – filme ale căror încasări au depășit 200 de milioane de dolari – algoritmul s-a descurcat ceva mai bine, identificând corect filmele de succes în 47,3% din cazuri. Dar investitori, păziți-vă! În jur de 10% dintre filmele pe care algoritmul le-a considerat a fi capabile să devină filme de succes au ajuns să aibă încasări sub 20 de milioane de dolari, ceea ce, după standardele hollywoodiene, este o sumă de bani demnă de toată mila.

Celelalte studii întreprinse de atunci au încercat să îmbunătățească aceste predicții, însă până în prezent nimeni nu a făcut vreun progres semnificativ. Toate dovezile trimit la o singură concluzie: până când nu ai informații despre reacția celor dintâi spectatori, popularitatea viitoare a unui film este, în mare măsură, imprevizibilă. Când vine vorba despre alegerea filmelor de succes dintr-o multitudine, Goldman avea dreptate: nimeni nu are habar de nimic.

Prezicerea popularității este, așadar, o chestiune complicată. Nu există o modalitate simplă de a separa ceea ce ne place de motivul pentru care ne place lucrul respectiv, iar în domeniul creativității asta este, mai degrabă, o problemă numai bună pentru algoritmi. Dacă nu poți folosi popularitatea pentru a-ți semnala ce este „bun", cum poți să măsori atunci calitatea?

Iată un lucru important: dacă vrem ca algoritmii să aibă autonomie în cadrul artelor – fie să creeze opere noi, fie să ne ofere informații semnificative despre arta pe care o creăm noi înșine – vom avea nevoie de o modalitate de măsurare a calității pentru a merge mai departe. Trebuie să existe un mod obiectiv de a îndruma algoritmul în direcția corectă, un fel de adevăr întemeiat, empiric indubitabil pe care acesta să-l aibă ca referință. Un fel de analogie valabilă în domeniul artei pentru: „acest grup de celule este purtător de cancer" sau „acuzatul a continuat să comită infracțiuni, a recidivat, după eliberarea pe cauțiune". În lipsa acesteia, progresul este dificil. Nu putem să proiectăm un algoritm sau să găsim un cântec „bun", dacă nu putem să definim ce înțelegem prin „bun".

Din nefericire, în încercarea de a găsi o unitate de măsură obiectivă pentru calitate, ne confruntăm cu o întrebare filosofică foarte controversată care datează încă de pe vremea lui Platon, o chestiune care se dezbate de mai bine de 2 000 de ani: cum poate fi apreciată valoarea estetică a artei?

Unii filosofi – precum Gottfried Leibniz – susțin că, dacă există obiecte pe care le considerăm cu toții frumoase (spre exemplu, *David* al lui Michelangelo sau *Lacrimosa* lui Mozart), atunci ar trebui să existe o esență definibilă și măsurabilă a frumuseții care face ca un obiect de artă să fie, în mod obiectiv, mai bun decât altul.

Pe de altă parte, se întâmplă destul de rar ca toată lumea să fie de aceeași părere. Alți filosofi, cum ar fi David Hume,

susțin că „frumusețea se află în ochiul privitorului", cum spunem noi, englezii, adică: „nu-i frumos ce e frumos, ci-i frumos ce-mi place mie". Gândiți-vă la opera lui Andy Warhol, de exemplu, care unora le oferă o experiență estetică intensă, pe când altora li se pare imposibil de distins, sub aspect artistic, de o cutie cu supă.

Alții, printre care se numără și Immanuel Kant, au spus că adevărul este pe undeva, pe la mijloc. Că aprecierea frumuseții nu este în totalitate subiectivă, dar nu poate fi nici cu totul obiectivă. Aprecierea este senzorială, emoțională și intelectuală totodată – și, foarte important, ea se poate modifica în timp, în funcție de starea de spirit a observatorului.

Există, cu siguranță, dovezi care vin în sprijinul acestei idei. Fanii lui Banksy își amintesc, poate, cum a aranjat acesta un stand în Central Park din New York în anul 2013, vânzând în anonimitate picturi originale pe pânză, vopsite cu spray în alb și negru, pe care a cerut 60 de dolari bucata. Standul era ascuns într-un șir de altele asemănătoare în care turiștilor li se vindeau mărunțișurile obișnuite, așa că prețul cerut trebuie să li se fi părut trecătorilor destul de mare. Se scurseseră câteva ore până când cineva a decis să cumpere un tablou. În ziua aceea Banksy a câștigat, în total, suma de 420 de dolari.[14] Un an mai târziu, la o casă de licitații din Londra, un cumpărător a apreciat că valoarea estetică a acelorași opere de artă era suficient de mare încât să fie tentat să cheltuiască suma de 68 000 de lire sterline (115 000 de dolari americani la cursul de atunci) pe o singură pictură.[15]

Este adevărat că nu oricui îi place Banksy (Charlie Brooker – cel care a creat serialul *Black Mirror** – l-a descris odată ca pe „un aiurit [a cărui] operă pare uimitor de deșteaptă pentru proști").[16] Așa că s-ar putea susține că istorioara aceasta este doar dovada faptului că opera lui Banksy nu are o calitate inerentă. Nu este decât o șmecherie la modă (și o dovadă

* Vezi nota de la p. 70 (n. red.).

a mentalității de turmă), care atrage niște prețuri atât de mari, că-ți vine să plângi. Dar aprecierea noastră estetică, atât de capricioasă, a fost urmărită și în raport cu forme de artă care sunt de o incontestabilă calitate.

Exemplul meu preferat este un experiment realizat de *Washington Post* în 2007.[17] Ziarul i-a cerut violonistului de faimă internațională Joshua Bell să mai adauge o reprezentație pe lista sa de concerte cu casa închisă. Înarmat cu vioara sa Stradivarius de 3,5 milioane de dolari, Bell s-a postat lângă o scară rulantă dintr-o stație de metrou din Washington, DC, într-o dimineață, în timpul orei de vârf, a așezat o pălărie pe jos, să strângă banii donați de trecători și a cântat timp de 43 de minute. Așa cum a scris *Washington Post*, acolo se afla unul dintre „cei mai buni interpreți de muzică clasică din lume, care cânta muzica cea mai distinsă care fusese compusă vreodată la una dintre cele mai scumpe viori create vreodată". Rezultatul? Șapte persoane s-au oprit să-l asculte câteva minute. Mai bine de 1 000 au trecut, pur și simplu, pe lângă el. La sfârșitul reprezentației, Bell strânsese în pălărie suma infimă de 32,17 dolari.

Ceea ce considerăm noi a fi „bun", se modifică de asemenea. Apetitul pentru anumite genuri de muzică clasică a fost deosebit de rezistent la trecerea timpului, însă nu se poate spune același lucru și despre alte forme de artă. Armand Leroi, profesor de biologie evoluționistă la Imperial College din Londra, a studiat evoluția muzicii pop și a descoperit, în cursul analizei, dovezi clare ale modificării preferințelor noastre față de aceasta:

> Populația posedă intrinsec un prag de plictiseală. Există o tensiune care se acumulează atunci când oamenii simt nevoia de ceva nou.[18]

Ca să ofer un exemplu, gândiți-vă la tobele electronice și sintetizatoarele care au fost la modă în muzica pop a anilor 1980 – atât de la modă, încât diversitatea muzicală din topuri mai-mai c-a dispărut. „Totul suna á la Madonna la începutul carierei sau ca un cântec compus de cei de la Duran Duran",

ne lămureşte Leroi. „Dar poate că vă spuneţi: «Bine! Am ajuns în punctul culminant al muzicii pop. El este chiar acolo. A fost găsită formula supremă a succesului!»" Numai că, bineînţeles, acest punct nu a fost găsit. La scurt timp, odată cu apariţia hip hop-ului, diversitatea muzicală din topurile muzicale a explodat din nou. Avea hip hop-ul ceva special, ceva care dusese la o schimbare? l-am întrebat eu pe Leroi. Acesta mi-a răspuns: „Nu cred. Ar fi putut fi orice altceva, dar s-a întâmplat să fie genul hip hop. La care consumatorul american a reacţionat, spunând: «Ei bine, iată ceva nou. Daţi-ne mai mult din ăsta.»"

Ideea este următoarea: chiar dacă există unele criterii obiective care fac ca o operă de artă să fie mai bună decât o alta, atât timp cât contextul joacă, pentru noi, un rol în aprecierea estetică a artei, nu este posibil să creezi o unitate de măsură concretă pentru calitatea estetică, una care să fie valabilă în toate locurile şi în toate timpurile. Indiferent de tehnicile statistice, de scamatoriile inteligenţei artificiale sau de algoritmii de învăţare automată pe care îi întrebuinţezi, încercarea de a folosi cifre pentru a cuprinde esenţa excelenţei artistice este asemenea încercării de a prinde fumul cu mâinile.

Un algoritm are, însă, nevoie de *ceva* pentru a merge mai departe. Aşadar, în momentul în care renunţi la popularitate şi la calitatea inerentă, îţi mai rămâne un singur lucru care să poată fi cuantificat: un indicator de similaritate cu orice a existat mai înainte.

Se pot face încă multe lucruri utilizând indicatorii de similaritate. Când vine vorba de construirea unui motor de recomandări, cum sunt cele de la Netflix şi Spotify, similaritatea este, fără îndoială, indicatorul ideal. Ambele companii au o anumită modalitate de a ajuta utilizatorii să descopere filme şi cântece noi, iar, dacă se abonează la serviciile lor, ambele susţin că vor prezice cu acurateţe preferinţele utilizatorilor. Nu-şi pot seta algoritmii pe ceea ce este la modă în acel moment, pentru că utilizatorii vor fi asediaţi de sugestii de genul Justin Bieber

și *Peppa Pig – The Movie* („Filmul cu Purcelușa Peppa"). Nu și-i pot seta nici pe vreun fel de substitut pentru calitate, cum ar fi recenziile, pentru că atunci paginile lor ar fi inundate de manifestări artistice la care adormi de-a-n picioarelea de plictiseală, atunci când oamenii nu-și doresc decât să-și scoată încălțările după o zi lungă de muncă și să se piardă într-un thriller de doi bani sau să se uite la ceva cu Ryan Gosling vreo două ore.

Similaritatea, în schimb, îi permite algoritmului să se concentreze strict pe preferințele unei persoane. Ce muzică ascultă, ce filme urmărește, la ce revine în mod repetat? Apoi, programul poate folosi IMDb, Wikipedia, bloguri de muzică sau articole din reviste pentru a extrage câteva cuvinte-cheie pentru fiecare cântec, artist sau film. Dacă procedează în același fel pentru întregul catalog, îi va fi ușor apoi să găsească și să recomande alte cântece și filme care au *tag*-uri, „etichete descriptive", asemănătoare. În plus, poate găsi și alți utilizatori cărora le-au plăcut filme sau cântece similare, să vadă ce alte cântece și filme le-au plăcut acestora și să le recomande primului utilizator.

În niciun moment Spotify sau Netflix nu încearcă să-ți ofere cântecul sau filmul perfecte. Perfecțiunea le interesează prea puțin. Spotify Discover nu promite că va căuta singura formație din lume menită să corespundă în întregime și negreșit cu gusturile și cu dispoziția ta actuală. Recomandările algoritmilor îți oferă doar cântecele și filmele care sunt suficient de bune încât să te împiedice să fii dezamăgit. Îți oferă o modalitate inofensivă de a-ți petrece timpul liber. Din când în când, vor găsi ceva care îți va plăcea la nebunie, dar asta este un fel de scamatorie din partea lor, o păcăleală similară aceleia folosite de șarlatanii care practică ghicitul în palmă și care pot descoperi multe lucruri despre o persoană nu din liniile desenate în palma acesteia, ci observându-i comportamentul, vestimentația etc. Nu este necesară decât

o reușită, din când în când, pentru a crea impresia că ai avut norocul să descoperi o muzică nouă. Motoarele de căutare nu simt nevoia să aibă dreptate de fiecare dată.

Similaritatea funcționează de minune pentru motoarele de căutare. Când le ceri, însă, algoritmilor să creeze artă în lipsa unui etalon clar pentru calitate, ei bine, abia atunci lucrurile încep să devină interesante. Poate un algoritm să fie creativ dacă singura sa noțiune despre artă ține de ceva ce s-a petrecut în trecut?

„ARTIȘTII BUNI COPIAZĂ, ARTIȘTII MARI FURĂ" – PABLO PICASSO

În octombrie 1997, la Universitatea din Oregon au venit niște spectatori pentru a asista la un concert neobișnuit. Pe scenă, în fața lor, era așezat un singur pian. Apoi a luat loc în fața sa pianista Winifred Kerner, care a început să cânte trei bucăți muzicale scurte, distincte.

Una a fost o compoziție pentru pian mai puțin cunoscută scrisă de maestrul barocului, Johann Sebastian Bach. A doua a fost compusă în stilul lui Bach de Steve Larson, profesor de muzică la acea universitate. Iar cea de-a treia a fost compusă de un algoritm conceput în mod intenționat astfel încât să imite stilul lui Bach.

După ce au ascultat toate cele trei piese, spectatorilor li s-a cerut să ghicească în mod exact cui îi aparținea fiecare piesă. Spre stupoarea lui Steve Larson, majoritatea a considerat că piesa lui era cea care fusese, în realitate, compusă de calculator. Iar în oftaturile unanime de consternare încântată, spectatorilor li s-a dezvăluit că muzica pe care o consideraseră ca aparținându-i lui Bach nu era decât creația unei mașini.

Larson nu a fost deloc mulțumit. Într-un interviu acordat ziarului *New York Times* la scurt timp după experiment, el a declarat:

> Admirația mea față de muzica lui Bach este profundă, cosmică. Am fost foarte tulburat când mi-am dat seama că oamenii puteau fi păcăliți de un program de calculator.

Nu era singurul surprins neplăcut. David Cope, bărbatul care concepuse remarcabilul algoritm care a stat la baza bucății muzicale compuse de calculator, a mai întâlnit această reacție, căci, după cum mi-a mărturisit:

> [La început] am jucat ceea ce eu numeam „jocul" cu persoanele". Iar când acestea greșeau, se înfuriau. Erau, oricum, supărate pe mine pentru că inventasem tot acel concept. Deoarece creativitatea este considerată o însușire specific umană.[19]

Aceasta fusese și opinia lui Douglas Hofstadter, profesor de științe cognitive și autorul a numeroase cărți, care organizase concertul de la bun început. Cu câțiva ani înainte, în 1979, când i se decernase Premiul Pulitzer pentru cartea sa *Gödel, Escher, Bach**, Hofstadter își exprimase ferm opinia în această privință:

> Muzica este un limbaj al emoțiilor și până când programele [de calculator] nu vor avea emoții tot atât de complexe ca ale noastre, este imposibil ca un program să compună ceva frumos... Să te gândești că am putea să-i cerem unui „tonomat" dinainte programat să născocească piese pe care le-ar fi putut scrie Bach este o subestimare grotească și rușinoasă a profunzimilor spiritului uman.[20]

După ce a ascultat, însă, produsul algoritmului lui Cope, EMI (acronimul pentru *Experiments in Musical Intelligence*, „Experimente asupra Inteligenței Muzicale"), Hofstadter a recunoscut că, poate, lucrurile nu erau chiar atât de simple. Iată ce a mărturisit el în zilele care au urmat după experimentul de la Universitatea din Oregon:

* Publicată în limba română în traducerea Corinei Tiron, cu titlul *Gödel, Escher, Bach. Brilianta ghirlandă eternă*, Editura Humanitas, București, 2015 (n. tr.).

> Sunt uimit și tulburat de creațiile lui EMI. Singura mea consolare în acest moment vine din faptul că mi-am dat seama că EMI nu generează stil de la sine, căci este dependent de imitarea compozitorilor precedenți. Dar nici asta nu este o consolare prea mare. Spre absoluta mea deznădejde, [poate că] muzica înseamnă mult mai puțin decât am crezut eu vreodată.[21]

Prin urmare, care este adevărul? Excelența estetică este apanajul activității umane? Sau un algoritm este capabil să creeze artă? Și dacă niște spectatori nu au reușit să deosebească muzica lui EMI de aceea a unui mare maestru înseamnă că mașina a dat dovadă de o autentică creativitate?

Să încercăm să abordăm, pe rând, aceste întrebări, începând cu ultima. Pentru formarea unei opinii pertinente și bine documentate, merită să ne oprim puțin pentru a înțelege cum funcționează algoritmul.* David Cope a fost destul de generos, încât să mă lămurească în această privință.

Primul pas în construirea unui algoritm a fost traducerea muzicii lui Bach în ceva care să poată fi înțeles de un calculator:

> Trebuie să introduci într-o bază de date cinci reprezentări ale unei singure note: timpul de aplicare, durata, înălțimea sunetului, volumul sonor și instrumentul.

Pentru fiecare notă din compozițiile lui Bach, Cope a trebuit să introducă în calculator, meticulos, de mână, aceste cinci reprezentări numerice și gândiți-vă că Bach a compus 371 de corale! Dar există multe alte piese, care însumează zeci de mii de note și pentru fiecare dintre acestea au fost introduse cele cinci cifre! A fost nevoie de un efort monumental din partea lui Cope:

* Aș dori, de asemenea, să vă recomand insistent să căutați măcar o parte din muzica lui Cope care există on-line. Cred că preferata mea este piesa orchestrală în stilul lui Vivaldi, care a fost postată pe YouTube la 12 august 2012: https://www.youtube.com/watch?v=2kuY3BrmTfQ (link accesat cel mai recent la 28 martie 2019).

Luni întregi nu am făcut altceva, zi de zi, decât să introduc numere în calculator. Dar trebuie să mărturisesc: eu pot să devin complet obsedat de ceva.

Apoi Cope a luat fiecare timp, fiecare bătaie, din muzica lui Bach și a studiat ce se întâmpla ulterior. După fiecare notă interpretată într-o corală de Bach, Cope a înregistrat nota imediat următoare. A așezat totul laolaltă într-un fel de dicționar – o bancă de date în care algoritmul putea să caute un singur acord și să găsească o listă exhaustivă a tuturor direcțiilor în care condeiul lui Bach trimisese muzica în continuare.

Judecând din acest punct de vedere, EMI seamănă oarecum cu algoritmii pentru generarea textului predictiv din *smartphone*-ul dumneavoastră. Bazându-se pe enunțurile pe care le-ați scris în trecut, telefonul creează un dicționar de cuvinte pe care este posibil să doriți să le scrieți în continuare și vi le oferă ca sugestie pe măsură ce vă compuneți textul.*

Etapa finală a fost să lase mașina de capul ei. Cope alimenta sistemul cu un acord de început și instruia algoritmul să-l caute în dicționarul său și să decidă ce să cânte după aceea, alegând un nou acord la întâmplare din acea listă. Apoi algoritmul repeta procedura – căutând fiecare acord ulterior în dicționar pentru a alege următoarele note pe care să le cânte.

Rezultatul l-a constituit o compoziție cu totul originală care sună ca și când ar fi fost compusă de Bach însuși.**

---

* Algoritmul de text predictiv poate fi, de asemenea, folosit pentru a „compune" un SMS în stilul dumneavoastră. Nu trebuie decât să deschizi o pagină nouă și să îi oferi algoritmului câteva cuvinte de pornire, cum ar fi „m-am născut". După aceea selectezi, în mod repetat, cuvintele care apar pe ecran. Iată ce mi-a ieșit mie (sincer!). Începe destul de bine, dar devine puțin forțat pe parcurs: „(m-am născut) ca să fiu un om bun și aș fi fericită să fiu cu tine o mulțime de persoane știu că nu-mi primești e-mail-urile și nu am deloc timp pentru așa ceva" (n. a.)

** Cope a mai descoperit și că trebuia să țină cont de un număr de alte acorduri, pe măsură ce algoritmul funcționa în continuare. Lungimea unei fraze muzicale și lungimea unei bucăți muzicale, de

Sau poate chiar *este* Bach. În orice caz, aceasta este viziunea lui Cope:

> Bach a compus toate acordurile. Este ca și când ai lua o bucată de brânză parmezan și ai da-o pe răzătoare, după care ai încerca să o reconstitui. Tot parmezan va rămâne.

Indiferent cui i se cuvine meritul final, un lucru este sigur. Oricât de frumos ar suna muzica creată de EMI, aceasta se bazează pe simpla recombinare a operei existente. Ea imită tiparele existente în muzica lui Bach, în loc să compună, efectiv, o piesă muzicală diferită.

Mai recent, au fost creați și alți algoritmi care produc o muzică plăcută sub aspect estetic, ceea ce este un pas înainte față de simpla recombinare. O abordare de mare succes a reprezentat-o folosirea algoritmilor genetici – un alt tip de învățare automată, care încearcă să exploateze modul în care funcționează selecția naturală. În definitiv, dacă privim păunii înseamnă că evoluția știe câte ceva despre crearea frumuseții.

Ideea este simplă: în interiorul acestor algoritmi, notele sunt prelucrate ca și când ar fi un fel de ADN al muzicii. Totul începe de la o populație inițială de „cântece" – fiecare fiind un amalgam de note, luate la întâmplare și puse laolaltă. După multe „generații", algoritmul se inspiră din cântece, găsind și răsplătind elementele „frumoase" din muzică pentru a da naștere, cu timpul, unor compoziții din ce în ce „mai bune". Spun „frumoase" și „mai bune" însă, așa cum deja știm, nu există nicio modalitate de a stabili ce înseamnă exact oricare din aceste cuvinte. Algoritmul poate să creeze poeme și picturi tot atât de bine ca și muzică, numai că nu dispune – încă! – pentru a continua, decât de un anumit grad de similitudine cu tot ce a fost creat mai înainte.

Iar uneori asta este tot ce-ți trebuie. Dacă ești în căutarea unei melodii de fundal pentru site-ul tău sau pentru un video

exemplu, s-au dovedit a corespunde cu o compoziție pe care ar fi compus-o însuși Bach (n. a.).

pe YouTube care să sune, în linii mari, ca o piesă folk, nu-ți pasă că ea seamănă cu cele mai bune cântece folk din trecut. De fapt, nu-ți dorești decât ceva care să nu încalce drepturile de autor, și să nu-ți dea bătaia de cap de a trebui să compui ceva tu însuți. Iar dacă asta este tot ce dorești să obții, există câteva companii care te pot ajuta. Firmele britanice Jukebox și AI Music oferă deja acest gen de servicii, utilizând algoritmi care sunt capabili să creeze muzică. O parte din aceasta se va dovedi utilă. O parte va fi (aproape) originală. O parte va fi chiar frumoasă. Algoritmii sunt, fără îndoială, imitatori excelenți, dar nu și inovatori foarte buni.

Nu afirm acest lucru pentru a le face un deserviciu algoritmilor. Cea mai mare parte a muzicii compuse de ființele umane nu este, nici ea, deosebit de inovatoare. Dacă ar fi să-l întrebăm pe Armand Leroi, biologul evoluționist care a studiat evoluția culturală a muzicii pop, suntem puțin cam idealiști în privința inventivității umane. Până și succesele deosebite din topurile muzicale, spune el, ar putea fi generate de o mașină. Iată, de exemplu, ce părere are despre cântecul „Happy" al lui Pharrell Williams (ceva îmi spune că nu este un fan al acestui artist!):

> *„Happy, happy, happy, I'm so happy!"** Zău așa! Are, să spunem, cam șase cuvinte în textul ăsta. Este un cântec cum nu se poate mai robotic, care descurajează chiar și dorința omenească cea mai înălțătoare și fundamentală de a asculta o melodie optimistă de vară. Este cântecul cel mai idiot și simplist care poate fi încă acceptat. Iar dacă *acesta* este nivelul – ei bine, atunci nu este prea dificil pentru un algoritm să-l creeze.

Leroi nu are o părere prea bună nici despre calitățile lirice ale lui Adele:

> Dacă ar fi să analizezi oricare dintre cântecele ei, nu ai găsi niciun sentiment care să nu poată fi creat de un generator de cântece triste.

* „Fericit, fericit, fericit,/ Sunt atât de fericit!" (engl.) (n. tr.).

S-ar putea să nu fiți de aceeași părere (nici eu nu sunt sigură că sunt), dar este, fără îndoială, un argument în favoarea faptului că totuși cea mai mare parte a creativității umane – cum sunt și produsele algoritmilor care „compun" – nu este decât o recombinare a unor idei preexistente. Așa cum spunea Mark Twain:

> Nu există ceva care să se numească „o idee nouă". Noi nu facem decât să luăm, pur și simplu, o mulțime de idei vechi și le așezăm într-un fel de caleidoscop mental. Le învârtim puțin, iar ele alcătuiesc combinații noi și interesante. Le învârtim în continuare și obținem combinații noi la nesfârșit; numai că ele sunt aceleași vechi fragmente de sticlă colorată care au fost utilizate de-a lungul veacurilor.[22]

Pe de altă parte, Cope are o definiție a creativității foarte simplă care înglobează cu ușurință ceea ce pot algoritmii să facă:

> Creativitatea înseamnă doar să găsești o asociere între două lucruri care, în mod obișnuit, nu par să aibă ceva în comun.

Poate că așa este. Dar eu, una, tot simt că, dacă EMI și algoritmii asemenea lui chiar *dau* dovadă de creativitate, atunci este o formă destul de firavă de creativitate. Muzica lor poate fi frumoasă, dar nu este profundă. Și oricât m-aș strădui, nu prea pot să scap de sentimentul că, dacă privim produsele acestor mașini drept artă, atunci rămânem cu o viziune despre lume destul de sărăcită cultural. Sub aspect cultural poate fi o „mâncare reconfortantă". Însă nu va deveni niciodată artă cu A mare.

Pe când mă documentam pentru scrierea acestui capitol, mi-am dat seama că motivul nemulțumirii pe care o resimțeam față de algoritmii care produc artă își are originea într-o întrebare diferită: adevărata problemă nu este dacă mașinăriile pot fi creative (pentru că pot fi), adevărata problemă este ce se înțelege, de fapt, prin artă.

Eu sunt matematiciană. Pot, cu toată încrederea, să fac schimb de informații despre rezultate fals-pozitive și adevăruri absolute referitoare la gradele de acuratețe și la statistici. În domeniul artistic, însă, aș prefera să țin cont de părerea lui Lev Tolstoi. Consider, ca și el, că adevărata artă are în centrul ei legăturile umane și comunicarea emoțiilor. Așa cum a spus Tolstoi:

> Arta nu este o abilitate meșteșugărească, ea înseamnă transmiterea sentimentelor resimțite de artist.[23]

Dacă sunteți de acord cu argumentul adus de el, atunci există un motiv pentru care calculatoarele nu vor putea niciodată să producă artă autentică. Un motiv foarte frumos a fost exprimat de Douglas Hofstadter cu ani înainte de a se fi întâlnit cu EMI:

> Un „program" care să producă muzică [...] ar trebui să colinde de unul singur în jurul lumii, croindu-și drum prin labirintul vieții și trăind fiecare clipă. Ar trebui să înțeleagă bucuria și singurătatea transmise de vânt într-o noapte friguroasă, să tânjească după mâna iubitei, să înțeleagă inaccesibilitatea unui oraș îndepărtat, frângerea inimii și revenirea ei la viață după o trecere în neființă. Ar trebui să fi cunoscut resemnarea și dezgustul de viață, suferința și disperarea, hotărârea și reușita, pietatea și venerarea. Ar trebui să fi amestecat sentimente opuse, precum nădejdea și teama, neliniștea și jubilarea, seninătatea și așteptarea. Din el ar trebui să facă parte un simț al grației, umorului, ritmului, o senzație de neprevăzut – și, bineînțeles, o minunată conștientizare a magiei unei proaspete creații. În acestea și numai în acestea se adăpostesc izvoarele înțelegerii în muzică.[24]

S-ar putea, prea bine, să greșesc. Poate că, dacă arta algoritmilor începe să semene a creație umană autentică (așa cum s-a întâmplat cu EMI), o vom aprecia totuși și îi vom adăuga propriul nostru înțeles. În definitiv, trecutul îndelungat

al muzicii pop contrafăcute pare să indice că oamenii pot avea o reacție emoțională la ceva care nu face decât să semene cu o legătură autentică. Și, poate, în momentul în care aceste opere de artă algoritmice vor deveni un loc comun, iar noi ne vom da seama că arta nu emană de la o ființă umană, această legătură unidirecțională nu ne va mai deranja. La urma-urmei, oamenii leagă relații pline de simțire cu obiecte care nu le întorc afecțiunea – mult-iubiții urși de pluș din copilărie sau păianjenii transformați în animale de companie, de exemplu.

Însă pentru mine, arta adevărată nu poate fi creată din întâmplare. Algoritmii nu pot să ajungă chiar oriunde. Nu chiar orice poate fi cuantificat. Printre toate lucrurile absolut impresionante, năucitoare, pe care mi le pot transmite informațiile și statisticile, nu se numără și cum este să fii om.

ÎNCHEIERE

Rahinah Ibrahim era o arhitectă cu patru copii şi un soţ care trăia în străinătate, muncea ca voluntar la un spital din localitate şi, pe deasupra, care mai şi lucra să-şi termine doctoratul la Universitatea Stanford. Ca şi când viaţa ei nu ar fi fost destul de ocupată, tocmai suferise o histerectomie de urgenţă şi – deşi între timp îşi revenise destul de bine – întâmpina încă probleme când trebuia să stea în picioare, fără sprijin, o perioadă oarecare, în lipsa medicamentelor. Cu toate acestea, când s-a apropiat cea de-a douăzeci şi opta International Conference on System Sciences*, în ianuarie 2005, şi-a rezervat un bilet de avion spre Hawaii şi s-a pregătit să-şi prezinte lucrarea cea mai recentă colegilor ei din mediul ştiinţific.[1]

Dimineaţa devreme, pe 2 ianuarie 2005, când Ibrahim a ajuns la aeroportul din San Francisco împreună cu fiica ei, s-a apropiat de ghişeu, şi-a prezentat actele şi a întrebat-o pe angajata aeroportului dacă o putea ajuta să-şi procure un scaun cu rotile. Membrii personalului nu au fost serviabili. Numele ei a apărut pe ecranul calculatorului strălucind intermitent, semnalând că ar figura pe lista federală a celor care au interdicţie de zbor, o bază de date pusă la punct după 11 septembrie 2001, pentru a-i împiedica pe suspecţii de terorism să călătorească.

* Conferinţă Internaţională anuală despre Ştiinţa Sistemelor (engl.) (n. tr.).

Fiica adolescentă a lui Ibrahim, rămasă singură şi năucită lângă ghişeu, a sunat un prieten de familie şi i-a spus că mama ei tocmai fusese luată cu forţa de acolo după ce o încătuşaseră. Între timp, Ibrahim fusese aşezată pe bancheta din spate a unei maşini de poliţie şi dusă la secţie. Au percheziţionat-o sub *hijab**, au refuzat să o lase să-şi ia medicamentele şi au închis-o într-o celulă. Două ore mai târziu, un agent al Homeland Security** a sosit cu documentele de eliberare şi i-a spus că fusese scoasă de pe lista celor cărora le era interzis să zboare. Ibrahim a ajuns la conferinţa din Hawaii, după care a luat avionul către Malaysia natală, să-şi viziteze familia.

Ibrahim fusese trecută pe lista celor care aveau interdicţie de zbor când un agent FBI bifase căsuţa greşită pe un formular. S-ar putea ca eroarea să se fi datorat unei confuzii între Jemaah Islamiyah, o organizaţie teroristă cunoscută pentru atacurile cu bombă din Bali, din 2002, şi Jemaah Islam, o organizaţie profesională malaysiană pentru persoane care studiază în străinătate. Ibrahim era membră a celei din urmă şi nu avusese niciodată vreo legătură cu cea dintâi. Nu era decât o greşeală, însă una care a avut consecinţe dramatice. Imediat ce eroarea fusese introdusă în sistemul automat, ea dobândise o aură de autoritate care o făcuse imună la orice recurs. Întâlnirea din San Francisco nu a însemnat însă şi finalul poveştii.

La întoarcerea din călătorie, două luni mai târziu, când revenea în SUA din Malaysia, Ibrahim a fost din nou oprită la aeroport. De data aceasta, soluţionarea nu a venit atât de rapid. Îi fusese anulată viza de şedere în SUA pe motiv că fusese suspectată de legături cu o organizaţia teroristă. Cu toate că era mama unui cetăţean american, că locuia în

* „Văl" (ar. în original), obiect vestimentar tradiţional al femeilor musulmane, voal care se poartă pe cap şi care lasă faţa vizibilă (n. tr.).

** United States Department of Homeland Security (Departamentul de Securitate Internă al Statelor Unite), agenţie guvernamentală a SUA similară ministerelor de interne din alte ţări (n. red.).

San Francisco și activa la una dintre universitățile cele mai prestigioase din țară, lui Ibrahim nu i se permitea să revină în Statele Unite. Urma să fie nevoită să lupte aproape zece ani pentru a obține câștig de cauză și a-și reabilita numele. Aproape zece ani în care i s-a interzis să pună piciorul pe sol american. Și toate acestea, din cauza unei erori umane și a unei mașini cu autoritate nelimitată.

OM PLUS MAȘINĂ

Nu se poate pune la îndoială impactul profund pozitiv pe care l-a avut automatizarea în viața noastră. Algoritmii pe care i-am construit până în prezent s-ar putea lăuda cu o listă de realizări uimitor de impresionantă. Ei ne pot ajuta să diagnosticăm cancerul mamar, să prindem ucigași în serie și să evităm prăbușirea avioanelor; tot ei ne oferă accesul gratuit și facil la întreaga bogăție a cunoașterii omenești și pun în legătură instantaneu oameni de pe tot globul într-o manieră la care strămoșii noștri nu puteau decât să viseze. Numai că, în insistența noastră de a automatiza, în graba noastră de a rezolva multe dintre problemele lumii, se pare că am schimbat o problemă cu alta. Algoritmii – oricât de utili și de impresionanți ar fi – ne-au lăsat un ghem tare încâlcit de complicații de rezolvat.

Oriunde privești – în sistemul judiciar, în sistemul de sănătate, în poliție, chiar și în domeniul cumpărăturilor online – există probleme de confidențialitate, scăpări, erori, probleme legate de responsabilitate și de transparență care nu vor dispărea foarte ușor. Dacă luăm în considerarea doar algoritmii care există în prezent, ne confruntăm cu probleme de echitate care ajung până la esența a cine suntem ca ființe umane, cum ne dorim să arate societatea în care trăim și până unde ne putem adapta la autoritatea iminentă a unei tehnologii lipsite de simțire.

Dar poate că tocmai așa și trebuie. Poate că greșim atunci când privim algoritmii ca și cum ar reprezenta cine știe ce autoritate.

În primul rând, reticența noastră de a pune sub semnul întrebării puterea pe care o are un algoritm a deschis ușa oamenilor care doresc să ne exploateze. Documentându-mă pentru această carte, am întâlnit tot felul de vânzători de iluzii dornici să speculeze mituri și să profite de credulitatea noastră. În pofida convingătoarelor dovezi științifice care susțin contrariul, există persoane care vând forțelor de poliție și guvernelor lumii algoritmi despre care pretind că ar „prezice" dacă cineva este terorist sau pedofil, exclusiv pe baza trăsăturilor faciale ale acelei persoane! Altele insistă că algoritmul lor poate să sugereze modificarea unei singure replici dintr-un scenariu care va reuși să ducă, ea singură, la încasări mult mai mari ale filmului.* Alte persoane afirmă cu îndrăzneală – și fără nicio urmă de sarcasm – că algoritmul lor poate să-ți găsească singura și adevărata iubire.**

* M-am întâlnit, de fapt, cu președintele acestei companii pentru a-i lua un interviu. L-am întrebat dacă și-a validat vreodată algoritmii, să vadă dacă aceștia se ridicau la nivelul așteptărilor, iar el s-a lansat într-o poveste lungă despre modul în care analiza efectuată de rețeaua neurală a făcut ca o vedetă importantă de la Hollywood să fie exclusă dintr-o serie de filme. Când am remarcat că asta însemna doar că oamenii credeau în algoritmul său și nu că acesta funcționa, el a răspuns: „Ei bine, doar nu facem un experiment științific" (n. a.).

** Există o șmecherie la care se poate apela pentru a descoperi algoritmii care nu-s buni de nimic. Mie îmi place să o numesc Testul Magiei. Ori de câte ori auziți ceva despre un algoritm, vedeți dacă puteți scoate oricare dintre sintagmele la modă, cum ar fi „învățare automată", „inteligență artificială (IA)", „rețea neurală", înlocuind-o cu următorul cuvânt: „magie". Ceea ce rezultă are încă sens din punct de vedere gramatical? Sau ceva și-a pierdut înțelesul? Dacă nu mai are noimă, eu mi-aș face griji că, pe undeva, ceva miroase a... știți ce. Fiindcă mă tem că nici măcar în viitorul foarte îndepărtat nu vom „eradica foametea din lume cu ajutorul «magiei»" și nici nu vom

Dar chiar şi algoritmii care fac ceea ce susţin că fac îşi folosesc adesea în mod incorect autoritatea. Cartea este plină de istorioare despre răul pe care-l pot face aceştia: „instrumentul bugetar" folosit pentru a reduce în mod arbitrar ajutorul financiar pentru persoanele cu dizabilităţi din statul Idaho. Algoritmii pentru depistarea infractorilor care aveau să recidiveze şi care, pe baza informaţiilor din trecutul istoric, sunt mai înclinaţi să sugereze un punctaj de risc ridicat pentru inculpaţii de culoare... Sistemul de depistare a afecţiunilor renale care sileşte milioane de oameni să împărtăşească informaţiile cele mai personale şi mai intime fără consimţământul sau ştirea lor... Algoritmul din supermarket, care îi răpeşte unei adolescente prilejul de a-şi anunţa tatăl că a rămas însărcinată... Lista cu subiecţi strategici, menită să ajute victimele atacurilor armate, dar care a fost folosită de poliţie ca o listă cu suspecţii buni de arestat... Exemple de inechitate se întâlnesc pretutindeni.

Cu toate acestea, sublinierea deficienţelor algoritmilor poate sugera că există o alternativă perfectă spre care să tindem. M-am gândit mult şi m-am străduit să găsesc fie şi un singur exemplu de algoritm absolut echitabil. Chiar şi aceia care arată bine la prima vedere (cum ar fi sistemul autopilot în cazul avioanelor sau reţelele neurale care diagnostichează cancerul) au probleme dacă îi iei la bani mărunţi. Aşa cum aţi citit în capitolul „Automobilele autonome", sistemul autopilot îi poate face pe cei care s-au antrenat prin intermediul sistemului automat să aibă un mare dezavantaj atunci când sunt la volan sau când manevrează ei înşişi manşa unui avion. Există chiar o anumită îngrijorare că algoritmii de depistare a tumorilor, pe care i-am întâlnit în capitolul „Medicina" şi care par atât de miraculoşi, nu funcţionează chiar impecabil pentru pacienţii din toate grupurile etnice. Însă exemplele de

„folosi «magia» pentru a scrie scenariul perfect", aşa cum nu reuşim să facem toate aceste lucruri nici cu ajutorul IA (n. a.).

sisteme perfect echitabile, corecte, nu prea abundă nici când *nu* sunt implicați algoritmii. Oriunde te-ai uita, orice domeniu ai studia, dacă sapi suficient de adânc în interiorul oricărui sistem, vei descoperi un fel sau altul de discriminare.

Imaginați-vă, așadar, pentru o clipă: ce-ar fi să acceptăm faptul că perfecțiunea nu există? Algoritmii chiar *vor face greșeli*. Algoritmii chiar *vor fi nedrepți*. Asta nu ar trebui să ne împiedice în niciun fel să încercăm să-i facem mai precişi şi mai puțin discriminatorii ori de câte ori ne stă în putință, dar, probabil, recunoscând că algoritmii nu sunt perfecți, așa cum nu sunt nici oamenii, s-ar diminua orice prezumție în privința autorității de care dispun.

Închipuiți-vă că, în loc să ne concentrăm atenția exclusiv asupra construirii unor algoritmi care să corespundă unui standard imposibil de corectitudine desăvârşită, i-am realiza astfel încât să le facilităm remedierea atunci când, inevitabil, vor da greș; că am dedica la fel de mult timp şi efort ca să ne asigurăm că sistemele automate vor fi tot atât de uşor de contestat, pe cât sunt de implementat. Poate că soluția este să construim algoritmi care să fie discutabili de la bun început. Ce ar fi dacă i-am concepe să asiste oamenii în luarea deciziilor, în loc să le dea instrucțiuni? Să fie transparenți în privința modului în care au ajuns la o anumită concluzie, în loc să ne aducă doar la cunoștință un anumit rezultat.

În opinia mea, algoritmii cei mai buni sunt cei care țin cont de ființa umană în fiecare stadiu. Cei care știu că avem obiceiul de a manifesta prea multă încredere în produsul unei mașini, recunoscându-şi, totodată, propriile defecte şi afişându-şi incertitudinea cu mândrie, punând-o în centrul atenției tuturor.

Aceasta era una dintre cele mai bune particularități ale calculatorului Watson al IBM, câştigătorul emisiunii-concurs *Jeopardy*. Cu toate că formatul concursului însemna că mașina trebuia să se oprească la un singur răspuns, algoritmul prezenta şi o serie de alternative pe care le luase în considerare pe parcurs,

alături de un punctaj care arăta cât de multă încredere avea în corectitudinea fiecărui răspuns. Dacă punctajele de calcularea a probabilității ca un infractor să recidiveze ar include ceva similar, poate că judecătorilor le-ar fi mai ușor să pună sub semnul întrebării informația oferită de algoritm. Iar dacă algoritmii de recunoaștere facială ar prezenta un număr de potriviri posibile în loc să aleagă o singură față, poate că identificarea greșită nu ar mai reprezenta o problemă atât de mare.

Aceeași caracteristică face să fie atât de eficiente rețelele neurale care scanează lamelele cu țesut mamar canceros. Algoritmul nu decide ce pacienți au tumori. El reduce marea varietate de celule, la doar câteva zone suspecte pe care să le verifice patologul. Algoritmul nu obosește niciodată, iar patologul rareori emite un diagnostic greșit. Algoritmul și omul lucrează în parteneriat, profitând fiecare de calitățile celuilalt și acceptându-și reciproc defectele.

Mai sunt și alte exemple – inclusiv din lumea șahului, unde am început această carte. De când a pierdut în fața lui Deep Blue, Garry Kasparov nu a întors spatele calculatoarelor. Chiar dimpotrivă. A devenit, în schimb, un mare susținător al ideii de „Centaur Chess", în care un jucător uman și un algoritm colaborează pentru a concura cu o altă echipă hibridă. Algoritmul evaluează consecințele posibile în cazul fiecărei mutări, reducând șansele jucătorului de a face o gafă, în timp ce ființa umană rămâne la conducerea jocului.

Iată cum descrie Kasparov acest lucru:

> Când jucăm asistați de calculator, ne putem concentra asupra planificării strategiei în loc să pierdem atât de mult timp făcând calcule. Creativitatea umană devine esențială în aceste condiții.[24]

Rezultă de aici un șah jucat la un nivel mai înalt decât oricând. Un joc tactic desăvârșit și strategii frumoase și pline de semnificație. Tot ce este cel mai bun din ambele lumi.

Acesta este viitorul pe care sper eu că voi ajunge să-l trăiesc. Unul în care algoritmii aroganți și dictatoriali, care umplu multe dintre aceste pagini, vor face parte din trecut. Unul în care vom înceta să privim mașinile ca pe niște stăpâni obiectivi și vom începe să le tratăm așa cum am face-o cu oricare altă sursă de putere. Punându-le deciziile sub semnul întrebării, cercetându-le motivele, recunoscându-ne emoțiile, cerând să știm cine urmează să obțină un anumit profit, considerându-le responsabile de propriile greșeli și refuzând să ne complăcem într-o anumită situație. Cred că aceasta este soluția pentru un viitor în care efectul global al algoritmilor să fie o forță benefică pentru societate. Și este corect ca această sarcină să apese pe umerii noștri. Pentru că un lucru este sigur: niciodată nu au fost oamenii mai importanți decât în era mașinilor, a algoritmilor.

NOTE

Introducere

1. Robert A. Caro, *The Power Broker: Robert Moses and the Fall of New York,* Bodley Head, Londra, 2015, p. 318.
2. Există vreo două eseuri nemaipomenite despre acest subiect care merită citite. Mai întâi, Langdon Winner, „Do artifacts have politics?", *Daedalus*, vol. 109, nr. 1/1980, pp. 121–136 (on-line la https://www.jstor.org/stable/20024652, site accesat cel mai recent la 20 februarie 2019), care conține exemplul cu podul lui Moses. Și o versiune mai modernă: Kate Crawford, „Can an algorithm be agonistic? Ten scenes from life in calculated publics", în *Science, Technology and Human Values*, vol. 41, nr. 1/2016, pp. 77–92.
3. *Scunthorpe Evening Telegraph*, numărul din 9 aprilie 1996.
4. Chukwuemeka Afigbo (@nke_ise) a postat pe Twitter un scurt videoclip despre acest efect. Merită să-l căutați, dacă nu l-ați văzut încă. Apare și pe YouTube, la adresa: https://www.youtube. com/watch?v=87QwWpzVy7I (accesată cel mai recent la 20 februarie 2019).
5. Interviu acordat de Mark Zuckerberg postului de televiziune CNN la 21 martie 2018: „I'm really sorry that this happened", care a fost urcat și pe *YouTube*, putând fi vizionat la adresa https://www.youtube.com/watch?v=G6DOhioBfyY (accesată cel mai recent la 20 februarie 2019).

CAPITOLUL 1 – *Puterea*

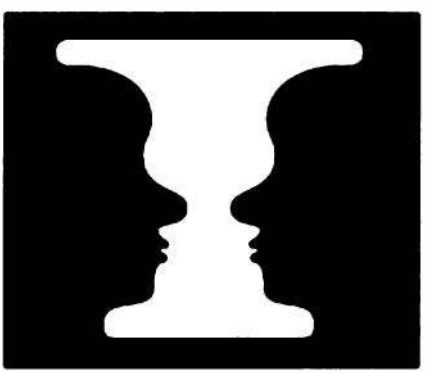
1. Informație obținută în timpul unei conversații particulare cu marele maestru de șah Jonathan Rowson.
2. Feng-Hsiung Hsu, „IBM's Deep Blue Chess grandmaster chips", *IEEE Micro*, vol. 19, nr. 2/1999, pp. 70–81, on-line la adresa: http://ieeexplore.ieee.org/document/755469/ (accesată cel mai recent la 20 februarie 2019).
3. Garry Kasparov, *Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins,* Hodder & Stoughton, Londra, 2017.
4. TheGoodKnight, „Deep Blue vs Garry Kasparov Game 2 (1997 Match)", YouTube, urcat pe 18 octombrie 2012, accesibil on-line la adresa https://www.youtube.com/watch?v=3Bd1Q2rOmok&t=2290s (accesată cel mai recent la 20 februarie 2019).
5. *Ibidem.*
6. Steven Levy, „Big Blue's Hand of God", *Newsweek*, numărul din 18 mai 1997, on-line la adresa http://www.newsweek.com/big-blues-hand-god-173076 (accesată cel mai recent la 20 februarie 2019).
7. Garry Kasparov, *op. cit.*, p. 187.
8. *Ibidem*, p. 191.
9. Definiția din *Merriam–Webster*. The *Oxford English Dictionary*'s subliniază bazele matematice ale algoritmilor: „un proces sau set de reguli care urmează să fie folosite în calcule matematice sau alte operații necesare în rezolvarea unei probleme oarecare, în principal, de către un computer".]
10. Există o multitudine de modalități diferite de grupare a algoritmilor și nu mă îndoiesc că unii specialiști în programare se vor plânge că lista este (mult prea) simplificată. E adevărat că o listă exhaustivă ar fi inclus și alte câteva categorii principale: algoritmi de mapare (*data mapping*, engl.), reducerea, regresia și gruparea sau clasificarea datelor (*clustering*, engl.; *le partitionnement de données,* fr.), ca să nu enumăr decât câțiva. În cele din urmă, am ales

acest tip anumit de categorii – apud Nicholas Diakopoulos, *Algorithmic Accountability Reporting: On the Investigation of Black Boxes,* Tow Center for Digital Journalism, Columbia University, New York, 2014, care acoperă foarte bine elementele de bază și oferă o modalitate utilă de a clarifica și de a condensa un domeniu de studiu vast și complex.

11. Kerbobotat, „Went to buy a baseball bat on Amazon, they have some interesting suggestions for accessories", *Reddit*, postare din 28 septembrie 2013, on-line la adresa https://www.reddit.com/r/funny/comments/1nb16l/went_to_buy_a_baseball_bat_on_amazon_they_have/ (accesată cel mai recent la 20 februarie 2019).
12. Sarah Perez, „Uber debuts a «smarter» UberPool in Manhattan", *TechCrunch*, 22 mai 2017, https://techcrunch.com/2017/05/22/uber-debuts-a-smarter-uberpool-in-manhattan/.
13. Intenționat am spus „teoretic". Realitatea ar putea fi ușor diferită. Unii algoritmi au fost concepuți, de-a lungul anilor, de către sute, chiar mii de programatori, fiecare adăugându-și, treptat, propriii pași în acest proces. Însă, pe măsură ce liniile de cod se înmulțesc, complexitatea sistemului crește, la rândul său, până când firele de execuție (*threads*, engl.) ajung ca o porție de spaghete încâlcite. În cele din urmă, algoritmul devine imposibil de urmărit și mult prea complicat pentru ca vreo ființă umană să-l mai poată înțelege. În anul 2013, companiei Toyota i s-a cerut să plătească suma de trei milioane de dolari ca despăgubire în urma unui accident mortal în care a fost implicat unul dintre autovehiculele lor. Acesta accelerase incontrolabil, deși conducătoarea auto apăsase pe frână, nicidecum pe accelerație în tot acel timp. Un specialist chemat să depună mărturie a declarat juriului că de vină era o instrucțiune accidentală ascunsă adânc în interiorul vastei încâlceli a software-ului. Vezi Phil Koopman, *A case study of Toyota unintended acceleration and software safety,* Pittsburgh: Carnegie Mellon University,

prezentare creată la 18 septembrie 2014, accesibilă online la adresa https://users.ece.cmu.edu/~koopman/pubs/koopman14_toyota_ua_slides.pdf (accesată cel mai recent la 20 februarie 2019).

14. Această iluzie optică (exemplul de aici este preluat de la adresahttps://commons.wikimedia.org/wiki/File:Vase_of_rubin.png, accesată cel mai recent la 20 februarie 2019) este cunoscută drept „Vaza lui Rubin", după Edgar Rubin, cel care a dezvoltat ideea. Este un exemplu de imagine *ambiguă* – situată chiar la limita dintre cele două profiluri închise la culoare care privesc unul spre celălalt și imaginea unei vaze albe. Așa cum apare desenată, este foarte ușor să treci de la una la alta din cele două în mintea ta și nu ai nevoie decât de două linii din imagine pentru a bascula într-o direcție și cealaltă. Poate de conturul vag al ochilor în cazul fețelor sau de umbra care mărginește gâtul vazei.



O chestiune asemănătoare este și exemplul de recunoaștere a imaginii câine/mașină. Echipa de cercetători a găsit o imagine care se situa chiar la limita dintre două clasificări diferite și a folosit cea mai mică perturbare posibilă pentru a trece imaginea dintr-o categorie într-alta, din perspectiva calculatorului.

15. Jiawei Su, Danilo Vasconcellos Vargas și Kouichi Sakurai, „One pixel attack for fooling deep neural networks", *arXiv:1719.08864v4* [cs.LG], publicat la 22 februarie 2018, https://arxiv.org/pdf/1710.08864.pdf (accesat cel mai recent la 20 februarie 2019).
16. Chris Brooke, „«I was only following satnav orders» is no defence: driver who ended up teetering on cliff edge convicted of careless driving", *Daily Mail*, numărul din 16 septembrie 2009, on-line la adresa http://www.dailymail.co.uk/news/article-1213891/Driver-ended-teetering-cliff-edge-guilty-blindly-following-sat-nav-directions.html#ixzz59vihbQ2n (accesată cel mai recent la 20 februarie 2019).
17. *Ibidem.*
18. Robert Epstein și Ronald E. Robertson, „The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections", *Proceedings of the National Academy of Sciences*, vol. 112, nr. 33/2015, pp. E4512–4521, on-line la adresa: http://www.pnas.org/content/112/33/E4512 (accesată cel mai recent la 20 februarie 2019).
19. David Shultz, „Internet search engines may be influencing elections", *Science*, numărul din 7 august 2015, on-line la adresa: http://www.sciencemag.org/news/2015/08/internet-search-engines-may-be-influencing-elections (accesată cel mai recent la 21 februarie 2019).
20. Robert Epstein și Ronald E. Robertson, *idem.*
21. Linda J. Cazul *KW versus Armstrong*, United States District Court for the District of Idaho, 2 mai 2012, proces a cărui hotărâre judecătorească poate fi citită on-line la adresa: https://scholar.google.co.uk/scholar_case?case=17062168494596747089&hl=en&as_sdt=2006 (accesată cel mai recent la 21 februarie 2019)Burdick, „Accountability and automation bias", *International Journal of Human–Computer Studies*, vol. 52/2000, pp. 701–717, on-line la adresa: http://lskitka.people.uic.edu/IJHCS2000.pdf (accesată cel mai recent la 21 februarie 2019).

22. Cazul *KW versus Armstrong*, United States District Court for the District of Idaho, 2 mai 2012, proces a cărui hotărâre judecătorească poate fi citită on-line la adresa: https://scholar.google.co.uk/scholar_case?case=17062168494596747089&hl=en&as_sdt=2006 (accesată cel mai recent la 21 februarie 2019).
23. Jay Stanley, *Pitfalls of Artificial Intelligence Decision making Highlighted in Idaho ACLU Case*, American Civil Liberties Union, 2 iunie 2017, on-line la adresa https://www.aclu.org/blog/privacy-technology/pitfalls-artificial-intelligence-decisionmaking-highlighted-idaho-aclu-case (accesată cel mai recent la 21 februarie 2019).
24. „K.W. v. Armstrong", *Leagle.com*, publicat la 24 martie 2014, la adresa: https://www.leagle.com/decision/infdco20140326c20 (accesată cel mai recent la 21 februarie 2019),
25. *Ibidem.*
26. ACLU Idaho staff., https://www.acluidaho.org/en/about/staff (site accesat cel mai recent la 21 februarie 2019).
27. Jay Stanley, *op. cit.*
28. ACLU, *Ruling mandates important protections for due process rights of Idahoans with developmental disabilities*, publicat la 30 martie 2016, on-line la adresa web: https://www.aclu.org/news/federal-court-rules-against-idaho-department-health-and-welfare-medicaid-class-action (accesată cel mai recent la 21 februarie 2019).
29. Jay Stanley, *op. cit.*
30. *Ibidem.*
31. *Ibid.*
32. *Ibid.*
33. *Ibid.*
34. Kristine Phillips, „The former Soviet officer who trusted his gut – and averted a global nuclear catastrophe", *Washington Post*, numărul din 18 septembrie 2017, on-line la adresa: https://www.washingtonpost.com/news/retropolis/wp/2017/09/18/

the-former-soviet-officer-who-trusted-his-gut-and-averted-a-global-nuclear-catastrophe/?utm_term=.6546e0f06cce (accesat cel mai recent la 21 februarie 2019).
35. Pavel Aksenov, „Stanislav Petrov: the man who may have saved the world", ştire şi articol BBC News, din 26 septembrie 2013, on-line la adresa http://www.bbc.co.uk/news/world-europe-24280831 (link accesat cel mai recent la 21 februarie 2019).
36. *Ibidem*.
37. Stephen Flanagan, *Re: Accident at Smiler Rollercoaster, Alton Towers, 2 June 2015: Expert's Report*, pregătit la cererea Direcţiei de Securitate şi Siguranţă, publicat on-line în octombrie 2015, la adresa http://www.chiark.greenend.org.uk/~ijackson/2016/Expert%20witness%20report%20from%20Steven%20Flanagan.pdf (accesată cel mai recent la 21 februarie 2019).
38. Paul E. Meehl, *Clinical versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence,* University of Minnesota, Minneapolis, 1996, (după prima ediţie din 1954), on-line la adresa http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.693.6031&rep=rep1&type=pdf (accesată cel mai recent la 21 februarie 2019).
39. William M. Grove, David H. Zald, Boyd S. Lebow, Beth E. Snitz şi Chad Nelson, „Clinical versus mechanical prediction: a meta-analysis", *Psychological Assessment*, vol. 12, nr. 1/2000, p. 19.
40. Berkeley J. Dietvorst, Joseph P. Simmons şi Cade Massey. „Algorithmic aversion: people erroneously avoid algorithms after seeing them err", *Journal of Experimental Psychology*, numărul din septembrie 2014, on-line la adresa: http://opim.wharton.upenn.edu/risk/library/WPAF201410-AlgorithmAversion-Dietvorst-Simmons-Massey.pdf (accesată cel mai recent la 21 februarie 2019).

CAPITOLUL 2 – *Datele personale informatizate*

1. Nicholas Carlson, „Well, these new Zuckerberg IMs won't help Facebook's privacy problems", *Business Insider*, numărul din 13 mai 2010, care poate fi citit on-line http://www.businessinsider.com/well-these-new-zuckerberg-ims-wont-help-facebooks-privacy-problems-2010-5?IR=T (link accesat cel mai recent la 6 martie 2019).
2. Clive Humby, Terry Hunt și Tim Phillips, *Scoring Points: How Tesco Continues to Win Customer Loyalty,* Kogan Page, Londra, 2008.
3. *Ibidem*, ediția Kindle, pp. 1313–1317.
4. See Eric Schmidt, „The creepy line", fragment de interviu urcat pe YouTube pe 11 februarie 2013, care poate fi vizionat on-line la adresa: https://www.youtube.com/watch?v=o-rvER6YTss (link accesat cel mai recent la 6 martie 2019).
5. Charles Duhigg, „How companies learn your secrets", *New York Times*, numărul din 16 februarie 2012, on-line la adresa: https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html (accesată cel mai recent la 6 martie 2019).
6. *Ibidem.*
7. Sarah Buhr, „Palantir has raised $880 million at a $20 billion valuation", *TechCrunch*, numărul din 23 decembrie 2015.
8. Comisia Federală pentru Comerț, *Data Brokers: A Call for Transparency and Accountability*, (Washington DC, mai 2014), on-line la adresa: https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (accesată cel mai recent la 6 martie 2019).
9. *Ibidem.*
10. Wolfie Christl, *Corporate Surveillance in Everyday Life*, raport publicat de Institute for Critical Digital Culture „Cracked Labs", Viena, iunie 2017, on-line la adresa http://crackedlabs.org/en/corporate-surveillance (accesată cel mai recent la 6 martie 2019).

11. Heidi Waterhouse, „The death of data: retention, rot, and risk", conferință organizată de The Lead Developer, Austin, Texas, susținută la 2 martie 2018, postată pe YouTube pe 21 martie 2018, on-line la adresa web: https://www.youtube.com/watch?v=mXvPChEo9iU (accesată cel mai recent la 6 martie 2019).
12. Amit Datta, Michael Carl Tschantz și Anupam Datta, „Automated experiments on ad privacy settings", *Proceedings on Privacy Enhancing Technologies*, nr. 1/2015, pp. 92–112.
13. Latanya Sweeney, „Discrimination in online ad delivery", în *Queue*, vol. 11, nr. 3/2013, p. 10, on-line la adresa web: https://dl.acm.org/citation.cfm?id=2460278 (accesată cel mai recent la 6 martie 2019).
14. Jon Brodkin, „Senate votes to let ISPs sell your Web browsing history to advertisers", în *Ars Technica*, numărul din 23 martie 2017, on-line la adresa web: https://arstechnica.com/tech-policy/2017/03/senate-votes-to-let-isps-sell-your-web-browsing-history-to-advertisers/ (accesată cel mai recent la 6 martie 2019).
15. Svea Eckert și Andreas Dewes, „Dark data", DEFCON, Conferința 25 din 20 octombrie 2017, postată pe YouTube pe 20 octombrie 2017, on-line la adresa web: https://www.youtube.com/watch?v=1nvYGi7-Lxo (accesată cel mai recent la 7 martie 2019).
16. Cercetătorii s-au bazat în această parte a activității lor pe lucrarea lui Arvind Narayanan și Vitaly Shmatikov, „Robust de-anonymization of large sparse datasets" prezentată în cadrul Simpozionului despre securitate și confidențialitate organizat de Institute of Electrical and Electronics Engineers (IEEE, Institutul Inginerilor Electrotehniști și Electroniști), în perioada 18–22 mai 2008.
17. Michal Kosinski, David Stillwell și Thore Graepel. „Private traits and attributes are predictable from digital records of human behavior", vol. 110, nr. 15/2013, pp. 5802–5805.
18. *Ibidem.*

19. Wu Youyou, Michal Kosinski și David Stillwell, „Computer-based personality judgements are more accurate than those made by humans", în *Proceedings of the National Academy of Sciences*, vol. 112, nr. 4/2015, pp. 1036–1040.
20. S.C. Matz, M. Kosinski, G. Nave și D.J. Stillwell, „Psychological targeting as an effective approach to digital mass persuasion", *Proceedings of the National Academy of Sciences*, vol. 114, nr. 48/2017, pp. 12714–12719, on-line la adresa https://www.pnas.org/content/114/48/12714?ijkey=67bf5df8e5307e45a2617411b054cbc508467668&keytype2=tf_ipsecsha (accesată cel mai recent pe 7 martie 2019).
21. Paul Lewis și Paul Hilder, „Leaked: Cambridge Analytica's blueprint for Trump victory", publicat în *Guardian*, numărul din 23 martie 2018, on-line la adresa https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory (accesată cel mai recent pe 7 martie 2019).
22. „Cambridge Analytica planted fake news", știre BBC din 20 martie 2018, on-line la adresa: http://www.bbc.co.uk/news/av/world-43472347/cambridge-analytica-planted-fake-news (accesată cel mai recent pe 7 martie 2019).
23. Adam D.I. Kramer, Jamie E. Guillory și Jeffrey T. Hancock, „Experimental evidence of massive-scale emotional contagion through social networks", *Proceedings of the National Academy of Sciences*, vol. 111, nr. 24/2014, pp. 8788–8790, on-line la adresa web: https://www.pnas.org/content/111/24/8788 (accesată cel mai recent pe 7 martie 2019).
24. Jamie Bartlett, „Big data is watching you – and it wants your vote", în *Spectator*, numărul din 24 martie 2018, on-line la adresa web: https://www.spectator.co.uk/2018/03/big-data-is-watching-you-and-it-wants-your-vote/ (accesată cel mai recent pe 7 martie 2019).
25. Li Xiaoxiao, „Ant Financial Subsidiary Starts Offering Individual Credit Scores", în *Caixin*, articol postat pe

2 martie 2015, on-line la adresa: https://www.caixinglobal.com/2015-03-02/101012655.html (accesată cel mai recent pe 7 martie 2019).

26. Rick Falkvinge, „In China, your credit score is now affected by your political opinions – and your friends' political opinions", în *Privacy News Online*, postat pe 3 octombrie 2015, on-line la adresa web: https://www.privateinternetaccess.com/blog/2015/10/in-china-your-credit-score-is-now-affected-by-your-political-opinions-and-your-friends-political-opinions/ (accesată cel mai recent pe 7 martie 2019).
27. Rogier Criemers, *State Council Guiding Opinions Concerning Establishing and Perfecting Incentives for Promise-keeping and Joint Punishment Systems for Trust-breaking, and Accelerating the Construction of Social Sincerity*, articol publicat pe blogul China Copyright and Media la 30 mai 2016, actualizat la 18 octombrie 2016, on-line la adresa web: https://chinacopyrightandmedia.wordpress.com/2016/05/30/state-council-guiding-opinions-concerning-establishing-and-perfecting-incentives-for-promise-keeping-and-joint-punishment-systems-for-trust-breaking-and-accelerating-the-construction-of-social-sincer/ (accesată cel mai recent la 7 martie 2019).
28. Rachel Botsman, *Who Can You Trust? How Technology Brought Us Together – and Why It Could Drive Us Apart*, Penguin, Londra, 2017, apud ediția Kindle, p. 151.

CAPITOLUL 3 – *Justiția*

1. John-Paul Ford Rojas, „London riots: Lidl water thief jailed for six months", articol în *Telegraph*, numărul din 7 ianuarie 2018, on-line la adresa http://www.telegraph.co.uk/news/uknews/crime/8695988/London-riots-Lidl-water-thief-jailed-for-six-months.html (accesară cel mai recent la 22 martie 2019).
2. Matthew Taylor, „London riots: how a peaceful festival in Brixton turned into a looting free-for-all", în *Guardian*,

numărul din 8 august 2011, on-line la adresa https://www.theguardian.com/uk/2011/aug/08/london-riots-festival-brixton-looting (accesată cel mai recent la 22 martie 2019).

3. John-Paul Ford Rojas, *op. cit.*
4. Josh Halliday, „London riots: how BlackBerry Messenger played a key role", în *Guardian*, numărul din 8 august 2011, on-line la adresa https://www.theguardian.com/media/2011/aug/08/london-riots-facebook-twitter-blackberry (accesată cel mai recent la 22 martie 2019).
5. David Mills, „Paul and Richard Johnson avoid prison over riots", în *News Shopper*, numărul din 13 ianuarie 2012, on-line la adresa http://www.newsshopper.co.uk/london-riots/9471288.Father_and_son_avoid_prison_over_riots/ (accesată cel mai recent la 22 martie 2019).
6. *Ibidem.*
7. John-Paul Ford Rojas, *op. cit.* „În mod normal, poliția nu te-ar aresta pentru acest gen de infracțiune. Nu te-ar ține în arest. Nu te-ar trimite în judecată" a scris, în 2015, Hannah Quirk, conferențiar universitar la catedra de justiție și drept penal a Universității din Manchester despre cazul lui Nicholas, apud. Carly Lightowlers și Hannah Quirk, „The 2011 English "riots": prosecutorial zeal and judicial abandon", în *British Journal of Criminology*, vol. 55, nr. 1/2015, pp. 65–85.
8. David Mills, *op. cit.*
9. William Austin și Thomas A. Williams III, „A survey of judges' responses to simulated legal cases: research note on sentencing disparity", în *Journal of Criminal Law and Criminology*, vol. 68, nr. 2/1977, pp. 306–310.
10. Mandeep K. Dhami și Peter Ayton, „Bailing and jailing the fast and frugal way", *Journal of Behavioral Decision-making*, în vol. 14, nr. 2/2001, pp. 141–68, on-line la adresa: http://onlinelibrary.wiley.com/doi/10.1002/bdm.371/abstract (accesată cel mai recent la 22 martie).
11. Aproape jumătate dintre judecători au avut opinii diferite despre cel mai bun curs al acțiunii în fiecare dintre cazuri.

12. Statisticienii au o metodă de măsurare a acestui gen de coerență a hotărârilor, numită „coeficientul lui Cohen" (kappa). Acesta ia în considerare faptul că (până și în cazul în care ai face o presupunere absolut la întâmplare) ai putea avea norocul să nimerești exact. Valoarea unu înseamnă o coerență perfectă. Valoarea zero e ca și cum ai fi ghicit la întâmplare. Valorile obținute de judecători au fost cuprinse între zero și unu, cu o medie de 0,69.
13. Diane Machin, „Sentencing guidelines around the world", o lucrare pregătită pentru Scottish Sentencing Council, în mai 2005, on-line la adresa: https://www.scottishsentencingcouncil.org.uk/media/1109/paper-31a-sentencing-guidelines-around-the-world.pdf (accesată cel mai recent la 22 martie 2019).
14. *Ibidem.*
15. *Ibid.*
16. Ernest W. Burgess, "Factors determining success or failure on parole", în *The Workings of the Intermediate-sentence Law and Parole System in Illinois* (State Board of Parole, Springfield, Illinois, 1928). Este o lucrare greu de găsit, așa că vă sugerez o lectură alternativă care îi aparține colegului lui Burgess, Tibbitts, despre un studiu ulterior celui original, „Success or failure on parole can be predicted: a study of the records of 3,000 youths paroled from the Illinois State Reformatory", în *Journal of Criminal Law and Criminology*, vol. 22, nr. 1/primăvara 1931, on-line la adresa https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=2211&context=jclc (accesată cel mai recent la 22 martie 2019). Celelalte categorii utilizate de Burgess erau „oaie neagră", „infractor din întâmplare", „greu de cap" și „bandit". El a descoperit că aceia din categoria „băiat de la fermă" erau cel mai puțin predispuși să recidiveze.
17. Karl F. Schuessler, „Parole prediction: its history and status", în *Journal of Criminal Law and Criminology*, vol. 45, nr. 4/1955, pp. 425–31, on-line la adresa: https://

pdfs.semanticscholar.org/4cd2/31dd25321a0c14a9358a93ebccb6f15d3169.pdf (accesată cel mai recent la 22 martie 2019).

18. *Ibidem*.
19. Bernard E. Harcourt, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age,* University of Chicago Press, Chicago și Londra, 2007, p. 1.
20. Philip Howard, Brian Francis, Keith Soothill și Les Humphreys, *OGRS 3: The Revised Offender Group Reconviction Scale*, Sinteza cercetării 7/09, Ministry of Justice (Ministerul de justiție britanic), Londra, 2009, on-line la adresa https://core.ac.uk/download/pdf/1556521.pdf (accesată cel mai recent la 22 martie 2019).
21. Se cuvine să fac un mic avertisment aici: în această statistică este posibil să existe o eroare de selecție. „Întreabă publicul" se folosea în mod obișnuit în primele etape ale jocului, când întrebările erau mult mai ușoare. Însă ideea că opiniile colective ale unui grup sunt mai corecte decât cele ale unei singure persoane, reprezintă un fenomen bine documentat. Pentru mai multe informații despre acest subiect, vezi James Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter than the Few,* Doubleday, New York, 2004, p. 4.
22. Netflix Technology Blog, on-line la adresa https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-2-d9b96aa399f5 (accesată cel mai recent la 22 martie 2019).
23. Shih-ho Cheng, „Unboxing the random forest classifier: the threshold distributions", Airbnb Engineering and Data Science, on-line la adresa: https://medium.com/airbnb-engineering/unboxing-the-random-forest-classifier-the-threshold-distributions-22ea2bb58ea6 (accesată cel mai recent la 22 martie 2019).
24. Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig și Sendhil Mullainathan, *Human Decisions and Machine Predictions*, Documentul de lucru nr. 23180 al

National Bureau of Economic Research (NBER, Biroul Național de Cercetări Economice, ONG american), National Bureau of Economic Research, Cambridge, Massachussetta, februarie 2017), on-line la adresa: http://www.nber.org/papers/w23180 (accesată cel mai recent la 22 martie 2019). Studiul acesta a folosit, de fapt „arbori de decizie de tip *gradient boosted*", un algoritm asemănător cu *random forests*. Ambii combină predicțiile unei multitudini de arbori de decizie pentru pentru a ajunge la un rezultat, însă arborii din metoda *gradient-boosted* cresc secvențial, pe când în *random forests* cresc în paralel. Pentru realizarea acestui studiu, setul de date a fost, mai întâi, înjumătățit. O jumătate a fost folosită pentru antrenarea algoritmului, cealaltă jumătate a fost păstrată deoparte. În momentul în care algoritmul a fost pregătit, a luat cazuri din jumătatea pe care nu o văzuse niciodată și a încercat să prevadă ce avea să se întâmple. (Dacă nu s-ar împărți, mai întâi, informațiile, algoritmul nu ar fi decât un frumos tabel de căutare.)

25. Oamenii de știință au petrecut foarte mult timp dezvoltând tehnici statistice care să se ocupe exact de această problemă, astfel încât să se poată face, în continuare, o comparație semnificativă între predicțiile făcute de judecători și, respectiv, de algoritmi. Pentru mai multe informații legate de această temă, vezi Jon Kleinberg, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig și Sendhil Mullainathan, *op. cit.*
26. „Costs per place and costs per prisoner by individual prison", *National Offender Management Service Annual Report and Accounts 2015–16*, Act adițional de gestionare a informațiilor din Comunicatul Ministerului de Justiție, 27 octombrie 2016, on-line la adresa: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/563326/costs-per-place-cost-per-prisoner-2015-16.pdf (accesată cel mai recent la 22 martie 2019).
27. Marc Santora, „City's annual cost per inmate is $168,000, study finds", în *New York Times*, 23 august 2013, on-line la

adresa: http://www.nytimes.com/2013/08/24/nyregion/citys-annual-cost-per-inmate-is-nearly-168000-study-says.html (accesată cel mai recent la 22 martie 2019); Universitatea Harvard, „Harvard at a glance", on-line la adresa https://www.harvard.edu/about-harvard/harvard-glance (accesată cel mai recent la 22 martie 2019).

28. Luke Dormehl, *The Formula: How Algorithms Solve All Our Problems... and Create More,* W.H. Allen, Londra, 2014, p. 123.
29. Julia Angwin, Jeff Larson, Surya Mattu și Lauren Kirchner, „Machine bias", în ProPublica, 23 mai 2016, on-line la adresa https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (accesată cel mai recent la 22 martie 2019).
30. Chestionar de „Evaluare a riscurilor", on-line la adresa https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html (accesată cel mai recent la 22 martie 2019).
31. Tim Brennan, William Dieterich și Beate Ehret (Institutul Northpointe), „Evaluating the predictive validity of the COMPAS risk and needs assessment system", în *Criminal Justice and Behavior*, vol. 36, nr. 1/2009, pp. 21–40, on-line la adresa http://www.northpointeinc.com/files/publications/Criminal-Justice-Behavior-COMPAS.pdf (accesată cel mai recent la 22 martie 2019). Conform unui studiu din 2018, algoritmul COMPAS are o precizie asemănătoare cu un „ansamblu" de persoane. Cercetătorii au demonstrat că, dacă cereau unui grup de 20 de persoane fără experiență să prezică recidiva, obțineau un punctaj echivalent cu acela al sistemului COMPAS. Este o comparație interesantă, însă merită să ne amintim că în sălile de judecată adevărate, judecătorii nu dispun de un grup de persoane necunoscute care să voteze în camera din spate. Sunt pe cont propriu. Iar aceasta este singura comparație care contează cu adevărat. Vezi Julia Dressel și

Hany Farid, „The accuracy, fairness, and limits of predicting recidivism", în *Science Advances*, vol. 4, nr. 1/2018.

32. *Christopher Drew Brooks* v. *Commonwealth*, Curtea de Apel a statului Virginia, Memorandum de acord al judecătorului Rudolph Bumgardner III, 28 ianuarie 2004, on-line la adresa https://law.justia.com/cases/virginia/court-of-appeals-unpublished/2004/2540023.html (accesată cel mai recent la 22 martie 2019).
33. „ACLU brief challenges constitutionality of Virginia's sex offender risk assessment guidelines", American Civil Liberties Union Virginia, 28 octombrie 2003, on-line la adresa https://acluva.org/en/press-releases/aclu-brief-challenges-constitutionality-virginias-sex-offender-risk-assessment (accesată cel mai recent la 22 martie 2019).
34. *State* v. *Loomis*, Curtea Supremă a statului Wisconsin,13 iulie 2016, on-line la adresa http://caselaw.findlaw.com/wi-supreme-court/1742124.html (accesată cel mai recent la 22 martie 2019).
35. Citatele provin dintr-o discuție particulară cu Richard Berk.
36. Julia Angwin, Jeff Larson, Surya Mattu și Lauren Kirchner, *op. cit*.
37. *Global Study on Homicide 2013,* Biroul ONU pentru Droguri și Criminalitate, Viena, 2014, on-line la adresa http://www.unodc.org/documents/gsh/pdfs/2014_GLOBAL_HOMICIDE_BOOK_web.pdf (accesată cel mai recent la 22 martie 2019).
38. ACLU, „The war on marijuana in black and white", iunie 2013, on-line la adresa www.aclu.org/files/assets/aclu-the waronmarijuana-ve12.pdf (accesată cel mai recent la 22 martie 2019).
39. În mod surprinzător, poate, poziția companiei Equivant asupra acestui aspect este susținută de Curtea Supremă a Statului Wisconsin. După ce Eric Loomis a fost condamnat la șase ani de închisoare de către un judecător care a utilizat instrumentul COMPAS de evaluare a riscului, el a

făcut apel împotriva acestei sentințe. În procesul *Loomis* vs. *Wisconsin*, s-a susținut că utilizarea unui software brevetat, cu circuit închis, de evaluare a riscului pentru stabilirea sentinței sale îi încălca dreptul la un proces echitabil, întrucât apărarea nu putea să pună sub semnul întrebării validitatea științifică a punctajului obținut. Însă Curtea Supremă a Statului Wisconsin a stabilit că utilizarea în instanță, cu titlu de încercare, a unei evaluări algoritmice a riscului în luarea deciziei nu încălca dreptul inculpatului la un proces echitabil: Curtea Supremă a Statului Wisconsin, cazul nr. 2015AP157-CR, verdict înregistrat la 13 iulie 2016, on-line la adresa https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690 (accesată cel mai recent la 22 martie 2019).

40. Lucy Ward, „Why are there so few female maths professors in universities?", *Guardian*, 11 martie 2013, on-line la adresa https://www.theguardian.com/lifeandstyle/the-womens-blog-with-jane-martinson/2013/mar/11/women-maths-professors-uk-universities (accesată cel mai recent la 22 martie 2019).
41. Sonja B. Starr și M. Marit Rehavi, *Racial Disparity in Federal Criminal Charging and Its Sentencing Consequences*, Program de Drept și Științe Economice, Document de lucru nr. 12-002, Universitatea de Drept din Statul Michigan, Ann Arbor, 7 mai 2012, on-line la adresa http://economics.ubc.ca/files/2013/05/pdf_paper_marit-rehavi-racial-disparity-federal-criminal.pdf (accesată cel mai recent la 22 martie 2019).
42. David Arnold, Will Dobbie și Crystal S. Yang, *Racial Bias in Bail Decisions*, NBER Document de lucru nr. 23421, Biroul Național de Cercetări Economice-NBER, Cambridge, Massachussetts, 2017, on-line la adresa https://www.princeton.edu/~wdobbie/files/racialbias.pdf (accesată cel mai recent la 22 martie 2019).
43. John J. Donohue III, *Capital Punishment in Connecticut, 1973–2007: A Comprehensive Evaluation from 4686*

Murders to One Execution, Şcoala de Drept Stanford şi Biroul Naţional de Cercetări Economice, Stanford, California şi Cambridge, Massachussetts, octombrie 2011, on-line la adresa https://law.stanford.edu/wp-content/uploads/sites/default/files/publication/259986/doc/slspublic/fulltext.pdf (accesată cel mai recent la 22 martie 2019).

44. Adam Benforado, *Unfair: The New Science of Criminal Injustice,* Crown, New York, 2015, p. 197.
45. Sonja B. Starr, *Estimating Gender Disparities in Federal Criminal Cases,* Universitatea Michigan, Facultatea de Drept şi Articol ştiinţific pe teme de economie nr. 12-018, Universitatea de Drept din Statul Michigan, Ann Arbor, 29 august 2012, on-line la adresa https://ssrn.com/abstract=2144002 ori la http://dx.doi.org/10.2139/ssrn.2144002 (accesată cel mai recent la 22 martie 2019).
46. David B. Mustard, „Racial, ethnic, and gender disparities in sentencing: evidence from the US federal courts", *Journal of Law and Economics,* vol. 44, nr. 2/aprilie 2001, pp. 285–314, on-line la adresa http://people.terry.uga.edu/mustard/sentencing.pdf (accesată cel mai recent la 22 martie 2019).
47. Daniel Kahneman, *Thinking, Fast and Slow,* Farrar, Straus & Giroux, New York, 2011, p. 44 [*Gândire rapidă, gândire lentă,* traducere în limba română de Dan Crăciun, Editura Publica, Bucureşti, 2015].
48. Chris Guthrie, Jeffrey J. Rachlinski şi Andrew J. Wistrich, *Blinking on the Bench: How Judges Decide Cases,* lucrarea nr. 917, Facultatea de Drept a Universităţii Cornell, New York, 2007, on-line la adresa http://scholarship.law.cornell.edu/facpub/917 (accesată cel mai recent la 22 martie 2019).
49. Daniel Kahneman, *op. cit.*, p. 13.
50. *Ibidem.*, p. 415.
51. Mandeep K. Dhami şi Peter Ayton, *op. cit.*

52. Brian Wansink, Robert J. Kent și Stephen J. Hoch, „An anchoring and adjustment model of purchase quantity decisions", *Journal of Marketing Research*, vol. 35, 1998, pp. 71–81, http://foodpsychology.cornell.edu/sites/default/files/unmanaged_files/Anchoring-JMR-1998.pdf (accesată cel mai recent la 22 martie 2019).
53. Mollie Marti și Roselle Wissler, „Be careful what you ask for: the effect of anchors on personal injury damages awards", *Journal of Experimental Psychology: Applied*, vol. 6, nr. 2/2000, pp. 91–103.
54. Birte Englich și Thomas Mussweiler, „Sentencing under uncertainty: anchoring effects in the courtroom", *Journal of Applied Social Psychology*, vol. 31, nr. 7/2001, pp. 1535–1551, on-line la adresa http://onlinelibrary.wiley.com/doi/10.1111/j.1559-1816.2001.tb02687.x (accesată cel mai recent la 22 martie 2019).
55. Birte Englich, Thomas Mussweiler și Fritz Strack, „Playing dice with criminal sentences: the influence of irrelevant anchors on experts' judicial decision making", *Personality and Social Psychology Bulletin*, vol. 32/2006, pp. 188–200, on-line la adresa https://www.researchgate.net/publication/7389517_Playing_Dice_With_Criminal_Sentences_The_Influence_of_Irrelevant_Anchors_on_Experts%27_Judicial_Decision_Making?enrichId=rgreqf2fedfeb71aa83f8fad80cc24df3254d-XXX&enrichSource=Y292ZXJQYWdlOzczODk1MTc7QVM6MTAzODIzNjIwMTgyMDIyQDE0MDE3NjQ4ODgzMTA%3D&el=1_x_3&_esc=publicationCoverPdf (accesată cel mai recent la 22 martie 2019).
56. *Ibidem*.
57. *Ibid*.
58. Mandeep K. Dhami, Ian K. Belton, Elizabeth Merrall, Andrew McGrath și Sheila Bird, „Sentencing in doses: is individualized justice a myth?", în curs de revizuire. Comunicată cu amabilitate într-o discuție particulară cu Mandeep Dhami.

59. *Ibidem.*
60. Adam N. Glynn și Maya Sen, „Identifying judicial empathy: does having daughters cause judges to rule for women's issues?", *American Journal of Political Science*, vol. 59, nr. 1/2015, pp. 37–54, on-line la adresa https://scholar.harvard.edu/files/msen/files/daughters.pdf (accesată cel mai recent la 22 martie 2019).
61. Shai Danziger, Jonathan Levav și Liora Avnaim-Pesso, „Extraneous factors in judicial decisions", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, nr. 17/2011, pp. 6889–6892, on-line la adresa http://www.pnas.org/content/108/17/6889 (accesată cel mai recent la 22 martie 2019).
62. Keren Weinshall-Margel și John Shapard, „Overlooked factors in the analysis of parole decisions", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, nr. 42,/2011, E833, on-line la adresa http://www.pnas.org/content/108/42/E833.long (accesată cel mai recent la 22 martie 2019).
63. Uri Simonsohn și Francesca Gino, „Daily horizons: evidence of narrow bracketing in judgment from 9,000 MBA-admission interviews", *Psychological Science*, vol. 24, nr. 2/2013, pp. 219–224, on-line la adresa https://ssrn.com/abstract=2070623 (accesată cel mai recent la 22 martie 2019).
64. Lawrence E. Williams și John A. Bargh, „Experiencing physical warmth promotes interpersonal warmth", *Science*, vol. 322, nr. 5901, pp. 606–607, on-line la adresa https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2737341/ (accesată cel mai recent la 22 martie 2019).

CAPITOLUL 4 – *Medicina*

1. Richard M. Levenson, Elizabeth A. Krupinski, Victor M. Navarro și Edward A. Wasserman. „Pigeons (*Columba livia*) as trainable observers of pathology and radiology breast cancer images", *PLOSOne*, 18 noiembrie 2015,

on-line la adresa http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0141357 (accesată cel mai recent la 22 martie 2019).
2. „Hippocrates' daughter as a dragon kills a knight, în „The Travels of Sir John Mandeville"", *British Library Online Gallery*, 26 martie 2009, on-line la adresa http://www.bl.uk/onlinegallery/onlineex/illmanus/harlmanucoll/h/011hrl000003954u00008v00.html (accesată cel mai recent la 22 martie 2019).
3. Eleni Tsiompanou, „Hippocrates: timeless still", JLL Bulletin: Commentaries on the History of Treatment Evaluation, James Lind Library, Oxford şi Edinburgh, 2012, on-line la adresa http://www.jameslindlibrary.org/articles/hippocrates-timeless-still/ (accesată cel mai recent la 22 martie 2019).
4. David K. Osborne, „Hippocrates: father of medicine", *Greek Medicine.net*, 2015, on-line la adresa http://www.greekmedicine.net/whos_who/Hippocrates.html (accesată cel mai recent la 22 martie 2019).
5. Richard Colgan, „Is there room for art in evidence-based medicine?", *AMA Journal of Ethics*, Virtual Mentor 13: 1, ianuarie 2011, pp. 52–54, on-line la adresa http://journalofethics.ama-assn.org/2011/01/msoc1-1101.html. (accesată cel mai recent la 22 martie 2019)
6. Joseph Needham, *Science and Civilization in China*, vol. 6, *Biology and Biological Technology*, partea a VI-a, *Medicine*, Nathan Sivin (editor), Cambridge University Press, Cambridge, 2004, p. 143, on-line la adresa https://monoskop.org/images/1/16/Needham_Joseph_Science_and_Civilisation_in_China_Vol_6-6_Biology_and_Biological_Technology_Medicine.pdf (accesată cel mai recent la 22 martie 2019).
7. „Ignaz Semmelweis", *Brought to Life: Exploring the History of Medicine,* Muzeul Ştiinţei, Londra, f.a., on-line la adresa http://broughttolife.sciencemuseum.org.uk/

broughttolife/people/ignazsemmelweis (accesată cel mai recent la 22 martie 2019).

8. Citatele provin dintr-o discuție particulară purtată cu Andy Beck.
9. Joann G. Elmore, Gary M. Longton, Patricia A. Carney, Berta M. Geller, Tracy Onega, Anna N.A. Tosteson, Heidi D. Nelson, Margaret S. Pepe, Kimberly H. Allison, Stuart J. Schnitt, Frances P. O'Malley și Donald L. Weaver, „Diagnostic concordance among pathologists interpreting breast biopsy specimens", *Journal of the American Medical Association*, vol. 313, nr. 11/17 martie 2015, 1122–1132, on-line la adresa https://jamanetwork.com/journals/jama/fullarticle/2203798 (accesată cel mai recent la 22 martie 2019).
10. *Ibidem*.
11. Numele de „rețea neurală" a apărut ca o analogie la ce se întâmplă în creier. Acolo, miliarde de neuroni sunt conectați între ei, într-o rețea gigantică. Fiecare neuron reacționează la propriile conexiuni și emite un semnal atunci când detectează starea de excitare a altui neuron. Semnalul duce apoi la excitarea altor neuroni care îl recepționează. O rețea neurală este o versiune mult simplificată și mult mai ordonată a creierului. Neuronii ei (artificiali) sunt structurați pe niveluri și toți neuronii de la un anumit nivel ascultă de toți neuronii de la nivelul anterior. În exemplul cu câinele, primul și cel dintâi nivel este constituit din pixelii individuali din imagine. Mai sunt apoi câteva niveluri având mii de neuroni în ele și un nivel final, care nu are decât un singur neuron de ieșire capabil să decidă dacă imaginea introdusă în rețea reprezintă un câine. Procedura de updatare a neuronilor este cunoscută drept „algoritm *backpropagation*". Începem cu neuronul final, de ieșire, care decide probabilitatea ca imaginea introdusă în rețea să reprezinte un câine. Să spunem că am introdus imaginea unui câine, iar el a prezis că acea imagine avea șanse de 70% să reprezinte

un câine. Neuronul de ieşire ia în considerare semnalele primite de la nivelul anterior şi spune: „Data viitoare când mai primesc o informaţie ca aceasta, voi majora probabilitatea ca imaginea să reprezinte un câine". Apoi îi comunică fiecărui neuron de pe nivelul anterior: „Hei, dacă mi-ai fi dat semnalul *acesta*, aş fi făcut o predicţie mai bună." Fiecare dintre acei neuroni îşi analizează semnalele de intrare şi îşi modifică semnalul de ieşire pentru data viitoare. Apoi comunică nivelului anterior ce semnale trebuia să fi trimis şi aşa mai departe prin toate nivelurile, până când se ajunge la cel dintâi. Acest proces de propagare înapoi a erorilor prin reţeaua neurală i-a dat numele de „algoritm *back*propagation" (de „propagare înapoi"). Pentru o perspectivă mai detaliată despre reţelele neurale, despre cum sunt construite şi antrenate acestea, vezi Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World,* Basic Books, New York, 2015.

12. Alex Krizhevsky, Ilya Sutskever şi Geoffrey E. Hinton, „ImageNet classification with deep convolutional neural networks", în F. Pereira, C.J.C. Burges, L. Bottou şi K.Q. Weinberger (editori), *Advances in Neural Information Processing Systems 25,* Neural Information Processing Systems Foundation, La Jolla, California, 2012, pp. 1097–1105, on-line la adresa http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf (accesată cel mai recent la 22 martie 2019). Acest tip de algoritmi este cunoscut sub numele de „reţea neurală de convoluţie". În loc de introducerea unei imagini întregi, algoritmul aplică, mai întâi, o serie de filtre şi caută tipare locale din modul în care este distorsionată imaginea.
13. Marco Tulio Ribeiro, Sameer Singh şi Carlos Guestrin, „«Why should I trust you?» Explaining the predictions of any classifier", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* San Francisco, 2016, pp. 1135–1144,

on-line la adresa http://www.kdd.org/kdd2016/papers/files/rfp0573-ribeiroA.pdf (accesată cel mai recent la 22 martie 2019).

14. Acest rezultat a fost comparat cu evaluarea făcută de mai mulți specialiști, a căror analiză colectivă a fost considerată ca reprezentând un „adevăr bazat pe observație, indubitabil empiric" în privința conținutului lamelelor.
15. Trafton Drew, Melissa L.H. Vo și Jeremy M. Wolfe, „The invisible gorilla strikes again: sustained inattentional blindness in expert observers", *Psychological Science*, vol. 24, nr. 9/septembrie 2013, pp. 1848–1853, on-line la adresa https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3964612/ (accesată cel mai recent la 22 martie 2019).
16. Gorila este poziționată în partea dreaptă sus a imaginii.
17. Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E. Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q. Nelson, Greg S. Corrado, Jason D. Hipp, Lily Peng și Martin C. Stumpe, „Detecting cancer metastases on gigapixel pathology images", Biblioteca Universității Cornell, 8 martie 2017, on-line la adresa https://arxiv.org/abs/1703.02442 (accesată cel mai recent la 22 martie 2019).
18. Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad și Andrew H. Beck, „Deep learning for identifying metastatic breast cancer", Biblioteca Universității Cornell, 18 iunie 2016, on-line https://arxiv.org/abs/1606.05718 (accesată cel mai recent la 22 martie 2019).
19. David A. Snowdon, „The Nun Study", *Boletin de LAZOS de la Asociación Alzheimer de Monterrey*, vol. 4, nr. 22/2000; D.A. Snowdon, „Healthy aging and dementia: findings from the Nun Study", *Annals of Internal Medicine*, vol. 139, nr. 5/septembrie 2003, pp. 450–454.
20. Densitatea ideilor – un substițut pentru complexitatea lingvistică – a fost calculată prin numărarea ideilor unice folosite de fiecare călugăriță într-un șir de zece cuvinte. Aveți o frumoasă prezentare aici: Associated

Press, „Study of nuns links early verbal skills to Alzheimer's", *Los Angeles Times*, 21 februarie 1996, on-line la adresa http://articles.latimes.com/1996-02-21/news/mn-38356_1_alzheimer-nuns-studied (accesată cel mai recent la 22 martie 2019).

21. Maja Nielsen, Jørn Jensen și Johan Andersen, „Pre-cancerous and cancerous breast lesions during lifetime and at autopsy: a study of 83 women", *Cancer*, vol. 54, nr. 4/1984, pp. 612–615, on-line la adresa http://onlinelibrary.wiley.com/wol1/doi/10.1002/1097-0142(1984)54:4%3C612::AID-CNCR2820540403%3E3.0.CO;2-B/abstract (accesată cel mai recent la 22 martie 2019).
22. H. Gilbert Welch și William C. Black, „Using autopsy series to estimate the disease "reservoir" for ductal carcinoma in situ of the breast: how much more breast cancer can we find?", *Annals of Internal Medicine*, vol. 127, nr. 11/decembrie 1997, pp. 1023–1028, on-line la adresa www.vaoutcomes.org/papers/Autopsy_Series.pdf (accesată cel mai recent la 22 martie 2019).
23. Obținerea unei statistici exacte este dificilă, pentru că ea depinde de țară și de demografie (precum și de intensitatea cu care o anumită țară face examinări pentru detectarea cancerului mamar). Pentru un rezumat de calitate, vezi: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer (accesată cel mai recent la 22 martie 2019).
24. Citatele provin dintr-o discuție particulară purtată cu Jonathan Kanevsky.
25. „Breakthrough method predicts risk of DCIS becoming invasive breast cancer", *Artemis*, mai 2010, on-line la adresa http://www.hopkinsbreastcenter.org/artemis/201005/3.html (accesată cel mai recent la 22 martie 2019).
26. H. Gilbert Welch, Philip C. Prorok, A. James O'Malley și Barnett S. Kramer, „Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness", *New England Journal of Medicine*, vol. 375/2016, pp. 1438–1447,

on-line la adresa http://www.nejm.org/doi/full/10.1056/NEJMoa1600249 (accesată cel mai recent la 22 martie 2019).
27. Un grup independent din Marea Britanie despre depistarea cancerului la sân, „The benefits and harms of breast cancer screening: an independent review", *Lancet*, vol. 380, nr. 9855/30 octombrie 2012, pp. 1778–1786, on-line la adresa http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(12)61611-0/abstract (accesată cel mai recent la 22 martie 2019).
28. Informații obținute în timpul unei discuții cu Jonathan Kanevsky.
29. Andrew H. Beck, Ankur R. Sangoi, Samuel Leung, Robert J. Marinelli, Torsten O. Nielsen, Marc J. van de Vijver, Robert B. West, Matt van de Rijn și Daphne Koller, „Systematic analysis of breast cancer morphology uncovers stromal features associated with survival", *Science Transitional Medicine*, 19 decembrie 2014, on-line la adresa https://becklab.hms.harvard.edu/files/becklab/files/sci_transl_med-2011-beck-108ra113.pdf (accesată cel mai recent la 22 martie 2019).
30. Phi Vu Tran, „A fully convolutional neural network for cardiac segmentation in short-axis MRI", 27 aprilie 2017, on-line la adresa https://arxiv.org/pdf/1604.00494.pdf (accesată cel mai recent la 22 martie 2019).
31. „Emphysema", *Imaging Analytics*, Zebra Medical, on-line la adresa https://www.zebra-med.com/algorithms/lungs/ (accesată cel mai recent la 22 martie 2019).
32. Eun-Jae Lee, Yong-Hwan Kim, Dong-Wha Kang et al., „Deep into the brain: artificial intelligence in stroke imaging", *Journal of Stroke*, vol. 19, nr. 3/2017, pp. 277–285, on-line la adresa https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5647643/ (accesată cel mai recent la 22 martie 2019).
33. Taylor Kubota, „Deep learning algorithm does as well as dermatologists in identifying skin cancer", *Stanford News*, 25 ianuarie 2017, on-line la adresa https://news.stanford.edu/2017/01/25/artificial-intelligence-used-identify-skin-cancer/ (accesată cel mai recent la 22 martie 2019).

34. Jo Best, „IBM Watson: the inside story of how the *Jeopardy*-winning supercomputer was born, and what it wants to do next", *Tech Republic*, f.a., on-line la adresa https://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next/ (accesată cel mai recent la 22 martie 2019).
35. Jennings Brown, „Why everyone is hating on IBM Watson, including the people who helped make it", *Gizmodo*, 14 august 2017, on-line la adresa https://www.gizmodo.com.au/2017/08/why-everyone-is-hating-on-watsonincluding-the-people-who-helped-make-it/ (accesată cel mai recent la 22 martie 2019).
36. Vezi https://www.theregister.co.uk/2017/02/20/watson_cancerbusting_trial_on_hold_after_damning_audit_report/ (accesată cel mai recent la 22 martie 2019).
37. Casey Ross și Ike Swetlitz, „IBM pitched its Watson supercomputer as a revolution in cancer care. It's nowhere close", *STAT*, 5 septembrie 2017, on-line la adresa https://www.statnews.com/2017/09/05/watson-ibm-cancer/ (accesată cel mai recent la 22 martie 2019).
38. Tomoko Otake, „Big data used for rapid diagnosis of rare leukemia case in Japan", *Japan Times*, 11 august 2016, on-line la adresa https://www.japantimes.co.jp/news/2016/08/11/national/science-health/ibm-big-data-used-for-rapid-diagnosis-of-rare-leukemia-case-in-japan/.Wf8S_hO0MQ8 (accesată cel mai recent la 22 martie 2019).
39. *Amyotrophic lateral sclerosis* (ALS) (engl.), „scleroză laterală amiotrofică". Vezi „Researchers validate five new genes responsible for ALS", *Science Daily*, 1 decembrie 2017, on-line la adresa https://www.sciencedaily.com/releases/2017/12/171201104101.htm (accesată cel mai recent la 22 martie 2019).
40. John Freedman, „A reality check for IBM's AI ambitions", *MIT Technology Review*, 27 iunie 2017.

41. *Asthma facts and statistics*, Asthma UK, 2016, on-line la adresa https://www.asthma.org.uk/about/media/facts-and-statistics/; *Asthma in the US*, Centers for Disease Control and Prevention, mai 2011, on-line la adresa https://www.cdc.gov/vitalsigns/asthma/index.html (accesată cel mai recent la 22 martie 2019).
42. „Schoolgirl, 13, who died of asthma attack was making regular trips to A&E and running out of medication – but was NEVER referred to a specialist even when her lips turned blue, mother tells inquest", *Daily Mail*, 13 octombrie 2015, on-line la adresa http://www.dailymail.co.uk/news/article-3270728/Schoolgirl-13-died-asthma-attack-not-referred-specialist-lips-turned-blue.html (accesată cel mai recent la 22 martie 2019).
43. *My Data, My Care: How Better Use of Data Improves Health and Wellbeing*, Richmond Group of Charities, Londra, ianuarie 2017, on-line la adresa https://richmondgroupofcharities.org.uk/publications (accesată cel mai recent la 22 martie 2019).
44. Terence Carney, „Regulation 28: report to prevent future deaths", raportul medico-legal în cazul Tamarei Mills, 29 octombrie 2015, on-line la adresa https://www.judiciary.gov.uk/publications/tamara-mills/ (accesată cel mai recent la 22 martie 2019).
45. Jamie Grierson şi Alex Hern, „Doctors using Snapchat to send patient scans to each other, panel finds", *Guardian*, 5 iulie 2017, on-line la adresa https://www.theguardian.com/technology/2017/jul/05/doctors-using-snapchat-to-send-patient-scans-to-each-other-panel-finds (accesată cel mai recent la 22 martie 2019).
46. Chiar dacă reuşeşti să rezolvi toate aceste chestiuni, uneori pur şi simplu nu există informaţii. Există mii de boli rare care au la bază cauze genetice şi care sunt efectiv unice. Medicii întâmpină dificultăţi imense în descoperirea vreunei astfel de boli, pentru că nu le-au mai întâlnit niciodată. Nici toţi algoritmii din lume nu pot rezolva

problema atunci când dispun de eșantioane extrem de reduse.

47. Hal Hodson, „Revealed: Google AI has access to huge haul of NHS patient data", *New Scientist*, 29 aprilie 2016, on-line la adresa https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/ (accesată cel mai recent la 22 martie 2019).
48. De fapt, cea mai mare parte a vinii pentru chestiunea așazisei „inadecvări juridice" a fost pusă în seama Royal Free Trust, care a fost puțin cam nerăbdător să se asocieze cu cea mai cunoscută companie de inteligență artificială din lume. Vezi scrisoarea lui Dame Fiona Caldicott, șefa protecției datelor naționale, care a ajuns la Sky News: Alex Martin, „Google received 1.6 million NHS patients' data on an «inappropriate legal basis»", Sky News, 15 mai 2017, on-line la adresa https://photos.google.com/share/AF1QipMdd5VTK0RNQ1AC3Dda1526CMG0vPD4P3x4x6_qmj0Zf101rbKyxfkfyputSPvqdA/photo/AF1QipP1_rnJMXkRyy3IuFHasilQHYEknKgnHFOFEy4T?key=U2pZUDM4bmo5RHhKYVptaDlkbEhfVFh4Rm1iVUVR (accesată cel mai recent la 22 martie 2019).
49. Denis Campbell, „Surgeons attack plans to delay treatment to obese patients and smokers", *Guardian*, 29 noiembrie 2016, on-line la adresa https://www.theguardian.com/society/2016/nov/29/surgeons-nhs-delay-treatment-obese-patients-smokers-york (accesată cel mai recent la 22 martie 2019).
50. Nir Eyal, „Denial of treatment to obese patients: the wrong policy on personal responsibility for health", *International Journal of Health Policy and Management*, vol. 1, nr. 2/ august 2013, pp. 107–110, on-line la adresa https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3937915/ (accesată cel mai recent la 22 martie 2019).
51. Pentru descrierea procedurilor, vezi http://galton.org/essays/1880-1889/galton-1884-jaigi-anthro-lab.pdf (accesată cel mai recent la 22 martie 2019).

52. Francis Galton, „On the Anthropometric Laboratory at the late international health exhibition", *Journal of the Anthropological Institute of Great Britain and Ireland*, vol. 14/1885, pp. 205–221.
53. „Taste", la adresa https://permalinks.23andme.com/pdf/samplereport_traits.pdf (accesată cel mai recent la 22 martie 2019).
54. „Sneezing on summer solstice?", *23andMeBlog*, 20 iunie 2012, on-line la adresa https://blog.23andme.com/health-traits/sneezing-on-summer-solstice/ (accesată cel mai recent la 22 martie 2019).
55. „Find out what your DNA says about your health, traits and ancestry", 23andMe, on-line la adresa https://www.23andme.com/en-gb/dna-health-ancestry/ (accesată cel mai recent la 22 martie 2019).
56. Kristen v. Brown, „23andMe is selling your data but not how you think", *Gizmodo*, 14 aprilie 2017, on-line la adresa https://gizmodo.com/23andme-is-selling-your-data-but-not-how-you-think-1794340474 (accesată cel mai recent la 22 martie 2019).
57. Michael Grothaus, „How23andMe is monetizing your DNA", *Fast Company*, 15 ianuarie 2015, on-line la adresa https://www.fastcompany.com/3040356/what-23andme-is-doing-with-all-that-dna (accesată cel mai recent la 22 martie 2019).
58. Rob Stein, „Found on the Web, with DNA: a boy's father", *Washington Post*, 13 noiembrie 2005, on-line la adresa http://www.washingtonpost.com/wp-dyn/content/article/2005/11/12/AR2005111200958.html (accesată cel mai recent la 22 martie 2019).
59. După un test ADN, tânărul a aflat că o anumită secvență repetitivă de pe cromozomul său Y – cel care se transmite de la tată la fiu – se mai întâlnea la două persoane purtând același nume de familie (rude îndepărtate din partea tatălui). Acel nume de familie, împreună cu locul și data

naşterii tatălui său au fost suficiente pentru a-i da acestuia de urmă.

60. M. Gymrek, A.L. McGuire, D. Golan, E. Halperin şi Y. Erlich, „Identifying personal genomes by surname inference", *Science*, vol. 339, nr. 6117/ianuarie 2013, pp. 321–324, on-line la adresa https://www.ncbi.nlm.nih.gov/pubmed/23329047 (accesată cel mai recent la 22 martie 2019).
61. În prezent, testele genetice pentru bolii Huntington nu sunt disponibile prin intermediul kit-urilor pentru analiza ADN-ului prezente pe piaţă.
62. Matthew Herper, „23andMe rides again: FDA clears genetic tests to predict disease risk", *Forbes*, 6 aprilie 2017, on-line la adresa https://www.forbes.com/sites/matthewherper/2017/04/06/23andme-rides-again-fda-clears-genetic-tests-to-predict-disease-risk/302aea624fdc (accesată cel mai recent la 22 martie 2019).

CAPITOLUL 5 – *Automobilele autonome*

1. DARPA, *Grand Challenge 2004: Final Report,* Defence Advanced Research Projects Agency, Arlington, 30 iulie 2004, on-line la adresa http://www.esd.whs.mil/Portals/54/Documents/FOID/Reading%20Room/DARPA/15-F-0059_GC_2004_FINAL_RPT_7-30-2004.pdf (accesată cel mai recent la 22 martie 2019).
2. *The Worldwide Guide to Movie Locations*, 7 septembrie 2014, on-line la adresa http://www.movie-locations.com/movies/k/Kill_Bill_Vol_2.html.WkYiqrTQoQ8 (accesată cel mai recent la 22 martie 2019).
3. Mariella Moon, *What you need to know about DARPA, the Pentagon's mad science division*, Engadget, 7 iulie 2014, on-line la adresa https://www.engadget.com/2014/07/07/darpa-explainer/ (accesată cel mai recent la 22 martie 2019).
4. DARPA, *Urban Challenge: Overview*, on-line la adresa http://archive.darpa.mil/grandchallenge/overview.html (accesată cel mai recent la 22 martie 2019).

5. Sebastian Thrun, „Winning the DARPA Grand Challenge, 2 August 2006", postat pe YouTube la 8 octombrie 2007, on-line la adresa https://www.youtube.com/watch?v=j8zj5lBpFTY (accesată cel mai recent la 22 martie 2019).
6. DARPA, *Urban Challenge: Overview, ed. cit.*
7. „DARPA Grand Challenge 2004 – road to...", prezentare postată pe YouTube la 22 ianuarie 2014, on-line la adresa https://www.youtube.com/watch?v=FaBJ5sPPmcI (accesată cel mai recent la 22 martie 2019).
8. Alex Davies, „An oral history of the DARPA Grand Challenge, the grueling robot race that launched the self-driving car", *Wired*, 8 martie 2017, on-line la adresa https://www.wired.com/story/darpa-grand-challenge-2004-oral-history/ (accesată cel mai recent la 22 martie 2019).
9. „Desert race too tough for robots", știre BBC News din 15 martie, 2004, on-line la adresa http://news.bbc.co.uk/1/hi/technology/3512270.stm (accesată cel mai recent la 22 martie 2019).
10. Alex Davies, *op. cit.*
11. Denise Chow, „DARPA and drone cars: how the US military spawned self-driving car revolution", *LiveScience*, 21 martie 2014, on-line la adresa https://www.livescience.com/44272-darpa-self-driving-car-revolution.html (accesată cel mai recent la 22 martie 2019).
12. Joseph Hooper, „From Darpa Grand Challenge 2004 DARPA's debacle in the desert", *Popular Science*, 4 iunie 2004, on-line la adresa https://www.popsci.com/scitech/article/2004-06/darpa-grand-challenge-2004darpas-debacle-desert (accesată cel mai recent la 22 martie 2019).
13. Alex Davies, *op. cit.*
14. DARPA, *Report to Congress: DARPA Prize Authority. Fiscal Year 2005 Report in Accordance with 10 U.S.C. 2374a*, martie 2006, on-line la adresa http://archive.darpa.mil/grandchallenge/docs/grand_challenge_2005_report_to_congress.pdf (accesată cel mai recent la 22 martie 2019).

15. Alan Ohnsman, „Bosch and Daimler to partner to get driverless taxis to market by early 2020s", *Forbes*, 4 aprilie 2017, on-line la adresa https://www.forbes.com/sites/alanohnsman/2017/04/04/bosch-and-daimler-partner-to-get-driverless-taxis-to-market-by-early-2020s/306ec7e63c4b (accesată cel mai recent la 22 martie 2019).
16. Ford, *Looking Further: Ford Will Have a Fully Autonomous Vehicle in Operation by 2021*, on-line la adresa https://corporate.ford.com/innovation/autonomous-2021.html (accesată cel mai recent la 22 martie 2019).
17. John Markoff, „Should your driverless car hit a pedestrian to save your life?", *New York Times*, 23 iunie 2016, on-line la addresa https://www.nytimes.com/2016/06/24/technology/should-your-driverless-car-hit-a-pedestrian-to-save-your-life.html (accesată cel mai recent la 22 martie 2019).
18. Clive Thompson, Anna Wiener, Ferris Jabr, Rahawa Haile, Geoff Manaugh, Jamie Lauren Keiles, Jennifer Kahn și Malia Wollan, „Full tilt: when 100 per cent of cars are autonomous", *New York Times*, 8 nov. 2017, https://www.nytimes.com/interactive/2017/11/08/magazine/tech-design-autonomous-future-cars-100-percent-augmented-reality-policing.htmlthe-end-of-roadkill (accesată cel mai recent la 22 martie 2019).
19. Peter Campbell, „Trucks headed for a driverless future: unions warn that millions of drivers' jobs will be disrupted", *Financial Times*, 31 ian. 2018, on-line la adresa https://www.ft.com/content/7686ea3e-e0dd-11e7-a0d4-0944c5f49e46 (accesată cel mai recent la 22 martie 2019).
20. Markus Maurer, J. Christian Gerdes, Barbara Lenz și Hermann Winner, *Autonomous Driving: Technical, Legal and Social Aspects,* Springer, New York, mai 2016, p 48.
21. Stephen Zavestoski și Julian Agyeman, *Incomplete Streets: Processes, Practices, and Possibilities,* Routledge, Londra, 2015, p. 29.

22. Markus Maurer, J. Christian Gerdes, Barbara Lenz şi Hermann Winner, *op. cit.* p. 53.
23. David Rooney, *Self-guided Cars,* Muzeul Ştiinţei, Lonfra, 27 aug. 2009, on-line la adresa https://blog.sciencemuseum.org.uk/self-guided-cars/ (accesată cel mai recent la 22 martie 2019).
24. Blake Z. Rong, „How Mercedes sees into the future", *Autoweek*, 22 ianuarie 2014, pn-line la adresa http://autoweek.com/article/car-news/how-mercedes-sees-future (accesată cel mai recent la 22 martie 2019).
25. Dean A. Pomerleau, *ALVINN: An Autonomous Land Vehicle In a Neural Network*, CMU-CS-89-107 (Pittsburgh: Carnegie Mellon University, ian. 1989), http://repository.cmu.edu/cgi/viewcontent.cgi?article=2874&context=compsci.
26. Joshua Davis, „Say hello to Stanley", *Wired*, 1 ianuarie 2006, on-line la adresa https://www.wired.com/2006/01/stanley/ (accesată cel mai recent la 22 martie 2019); iar pentru mai multe detalii, Dean A. Pomerleau, *Neural Network Perception for Mobile Robot Guidance,* Springer, New York, 2012, p. 52.
27. A. Filgueira, H. González-Jorge, S. Lagüela, L. Diaz-Vilariño şi P. Arias, „Quantifying the influence of rain in LiDAR performance", *Measurement*, vol. 95/ianuarie 2017, pp. 143–148, DOI: https://doi.org/10.1016/j.measurement.2016.10.009; https://www.sciencedirect.com/science/article/pii/S0263224116305577 (accesată cel mai recent la 22 martie 2019).
28. Chris Williams, „Stop lights, sunsets, junctions are tough work for Google's robo-cars", *The Register*, 24 august 2016, on-line la adresa https://www.theregister.co.uk/2016/08/24/google_self_driving_car_problems/ (accesată cel mai recent la 22 martie 2019).
29. Novatel, *IMU Errors and Their Effects*, on-line la adresa https://www.novatel.com/assets/Documents/Bulletins/APN064.pdf (accesată cel mai recent la 22 martie 2019).

30. Teorema în sine nu este decât o ecuație care face legătura între probabilitatea unei *ipoteze* dată fiind observarea unor indicii, și probabilitatea existenței acelor indicii dată fiind ipoteza. O expunere introductivă mai cuprinzătoare poate fi găsită on-line la adresa https://arbital.com/p/bayes_rule/?l=1zq (accesată cel mai recent la 22 martie 2019).
31. Sharon Bertsch McGrayne, *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy,* Yale University Press, New Haven, 2011.
32. M. Bayes și M. Price, *An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F.R.S. Communicated by Mr. Price, in a Letter to John Canton, A.M.F.R.S.* (1763), copie digitală încărcată în archive.org la 2 august 2011, https://archive.org/details/philtrans09948070 (accesată cel mai recent la 22 martie 2019).
33. Michael Taylor, „Self-driving Mercedes-Benzes will prioritize occupant safety over pedestrians", *Car and Driver*, 7 octombrie 2016, on-line la adresa https://blog.caranddriver.com/self-driving-mercedes-will-prioritize-occupant-safety-over-pedestrians/ (accesată cel mai recent la 22 martie 2019).
34. Jason Kottke, *Mercedes' Solution to the Trolley Problem*, Kottke.org, 24 octombrie 2016, on-line la adresa https://kottke.org/16/10/mercedes-solution-to-the-trolley-problem (accesată cel mai recent la 22 martie 2019).
35. Jean-François Bonnefon, Azim Shariff și Iyad Rahwan (2016), „The social dilemma of autonomous vehicles", *Science*, vol. 35/24 iunie 2016, on-line la adresa DOI 10.1126/science.aaf2654; https://arxiv.org/pdf/1510.03346.pdf (accesată cel mai recent la 22 martie 2019).

36. Toate citatele provin dintr-o discuție particulară purtată cu Paul Newman.
37. Naaman Zhou, „Volvo admits its self-driving cars are confused by kangaroos", *Guardian*, 1 iulie 2017, on-line la adresa https://www.theguardian.com/technology/2017/jul/01/volvo-admits-its-self-driving-cars-are-confused-by-kangaroos (accesată cel mai recent la 22 martie 2019).
38. Toate citatele provin dintr-o discuție particulară purtată cu Jack Stilgoe.
39. Jeff Sabatini, „The one simple reason nobody is talking realistically about driverless cars", *Car and Driver*, octombrie 2017, https://www.caranddriver.com/features/the-one-reason-nobody-is-talking-realistically-about-driverless-cars-feature (accesată cel mai recent la 22 martie 2019).
40. William Langewiesche, „The human factor", *Vanity Fair*, 17 septembrie 2014, on-line la adresa https://www.vanityfair.com/news/business/2014/10/air-france-flight-447-crash (accesată cel mai recent la 22 martie 2019).
41. Bureau d'Enquêtes et d'Analyses pour la Sécuritié de l'Aviation Civile, *Final Report on the Accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France Flight AF447 Rio de Janeiro – Paris*, ediția în limba engleză, Paris, actualizată în iulie 2012, on-line la adresa https://www.bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf (accesată cel mai recent la 22 martie 2019).
42. *Ibidem.*
43. William Langewiesche, *op. cit.*
44. *Ibidem.*
45. Jeff Wise, „What really happened aboard Air France 447", *Popular Mechanics*, 6 decembrie 2011, on-line la adresa http://www.popularmechanics.com/flight/a3115/what-really-happened-aboard-air-france-447-6611877/ (accesată cel mai recent la 22 martie 2019).
46. William Langewiesche, *op. cit.*

47. Jeff Wise, *op. cit.*
48. Lisanne Bainbridge, „Ironies of automation", *Automatica*, vol. 19, nr. 6/noiembrie 1983, pp. 775–9, on-line la adresa https://www.sciencedirect.com/science/article/pii/0005109883900468 (accesată cel mai recent la 22 martie 2019).
49. *Ibidem.*
50. Alex Davies, „Everyone wants a level 5 self-driving car – here's what that means", *Wired*, 26 iulie 2016.
51. Justin Hughes, „Car autonomy levels explained", *The Drive*, 3 noiembrie 2017, on-line la adresa http://www.thedrive.com/sheetmetal/15724/what-are-these-levels-of-autonomy-anyway (accesată cel mai recent la 22 martie 2019).
52. Lisanne Bainbridge, *op. cit.*
53. Jack Stilgoe, „Machine learning, social learning and the governance of self-driving cars", *Social Studies of Science*, vol. 48, nr. 1/2017, pp. 25–56.
54. Eric Tingwall, „Where are autonomous cars right now? Four systems tested", *Car and Driver*, octombrie 2017, on-line la adresa https://www.caranddriver.com/features/where-are-autonomous-cars-right-now-four-systems-tested-feature (accesată cel mai recent la 22 martie 2019).
55. Tracey Lindeman, „Using an orange to fool Tesla's autopilot is probably a really bad idea", *Motherboard*, 16 ianuarie 2018, on-line la adresa https://motherboard.vice.com/en_us/article/a3na9p/tesla-autosteer-orange-hack (accesată cel mai recent la 22 martie 2019).
56. Daisuke Wakabayashi, „Uber's self-driving cars were struggling before Arizona Crash", *New York Times*, 23 martie 2018, on-line la adresa https://www.nytimes.com/2018/03/23/technology/uber-self-driving-cars-arizona.html (accesată cel mai recent la 22 martie 2019).
57. Sam Levin, „Video released of Uber self-driving crash that killed woman in Arizona", *Guardian*, 22 martie 2018, on-line la adresa https://www.theguardian.com/

technology/2018/mar/22/video-released-of-uber-self-driving-crash-that-killed-woman-in-arizona (accesată cel mai recent la 22 martie 2019).
58. Audi, *The Audi vision of autonomous driving*, Audi Newsroom, 11 sept. 2017, on-line la adresa https://media.audiusa.com/en-us/releases/184 (accesată cel mai recent la 22 martie 2019).
59. P. Morgan, C. Alford și G. Parkhurst, *Handover Issues in Autonomous Driving: A Literature Review. Project Report*, University of the West of England, Bristol, iunie 2016, on-line la adresa http://eprints.uwe.ac.uk/29167/1/Venturer_WP5.2Lit%20ReviewHandover.pdf (accesată cel mai recent la 22 martie 2019).
60. William Langewiesche, *op. cit.*
61. Evan Ackerman, „Toyota's Gill Pratt on self-driving cars and the reality of full autonomy", *IEEE Spectrum*, 23 ianuarie 2017, on-line la adresa https://spectrum.ieee.org/cars-that-think/transportation/self-driving/toyota-gill-pratt-on-the-reality-of-full-autonomy (accesată cel mai recent la 22 martie 2019).
62. Julia Pyper, „Self-driving cars could cut greenhouse gas pollution", *Scientific American*, 15 septembrie 2014, on-line la adresa https://www.scientificamerican.com/article/self-driving-cars-could-cut-greenhouse-gas-pollution/ (accesată cel mai recent la 22 martie 2019).
63. Raphael E. Stern et al., „Dissipation of stop-and-go waves via control of autonomous vehicles: field experiments", *arXiv: 1705.01693v1*, 4 mai 2017, on-line la adresa https://arxiv.org/abs/1705.01693 (accesată cel mai recent la 22 martie 2019).
64. SomeJoe7777, „Tesla Model S forward collision warning saves the day", postare încărcată pe YouTube la 19 octombrie 2016, on-line la adresa https://www.youtube.com/watch?v=SnRp56XjV_M (accesată cel mai recent la 22 martie 2019).

65. Jordan Golson şi Dieter Bohn, „All new Tesla cars now have hardware for "full self-driving capabilities": but some safety features will be disabled initially", *The Verge*, 19 octombrie 2016, on-line la adresa https://www.theverge.com/2016/10/19/13340938/tesla-autopilot-update-model-3-elon-musk-update (accesată cel mai recent la 22 martie 2019).
66. Fred Lambert, „Tesla introduces first phase of "Enhanced Autopilot": "measured and cautious for next several hundred million miles" – release notes", *Electrek*, 1 ianuarie 2017, on-line la adresa https://electrek.co/2017/01/01/tesla-enhanced-autopilot-release-notes/ (accesată cel mai recent la 22 martie 2019).
67. DPC Cars, „Toyota Guardian and Chauffeur autonomous vehicle platform", postare încărcată pe YouTube la 27 septembrie 2017, on-line la adresa https://www.youtube.com/watch?v=IMdceKGJ9Oc (accesată cel mai recent la 22 martie 2019).
68. Brian Milligan, „The most significant development since the safety belt", ştire BBC News din 15 aprilie 2018, on-line la adresa http://www.bbc.co.uk/news/business-43752226(accesată cel mai recent la 22 martie 2019).

CAPITOLUL 6 – *Infracţionalitatea*

1. Bob Taylor, *Crimebuster: Inside the Minds of Britain's Most Evil Criminals*, Piatkus, Londra, 2002, capitolul 9, „A day out from jail".
2. *Ibidem.*
3. Nick Davies, „Dangerous, in prison – but free to rape", *Guardian*, 5 octombrie 1999, on-line la adresa https://www.theguardian.com/uk/1999/oct/05/nickdavies1 (accesată cel mai recent la 22 martie 2019).
4. João Medeiros, „How geographic profiling helps find serial criminals", *Wired*, noiembrie 2014, on-line la adresa http://www.wired.co.uk/article/mapping-murder (accesată cel mai recent la 22 martie 2019).

5. Nicole H. Rafter, ed., *The Origins of Criminology: A Reader,* Routledge, Abingdon, 2009, p. 271.
6. Luke Dormehl, *The Formula: How Algorithms Solve All Our Problems... and Create More,* W.H. Allen, Londra, 2014, p. 117.
7. *Ibidem*, p. 116.
8. D. Kim Rossmo, „Geographic profiling", în Gerben Bruinsma și David Weisburd (editori), *Encyclopedia of Criminology and Criminal Justice,* Springer, New York, 2014, on-line la adresahttps://link.springer.com/referencework entry/10.1007%2F978-1-4614-5690-2_678.
9. *Ibidem.*
10. João Medeiros, *op. cit.*
11. *Ibidem.*
12. „Sadistic" serial rapist sentenced to eight life terms", *Independent* (Irlanda), 6 octombrie 1999, on-line la adresa http://www.independent.ie/world-news/sadistic-serial-rapist-sentenced-to-eight-life-terms-26134260.html (accesată cel mai recent la 22 martie 2019).
13. *Ibidem.*
14. Steven C. Le Comber, D. Kim Rossmo, Ali N. Hassan, Douglas O. Fuller și John C. Beier, „Geographic profiling as a novel spatial tool for targeting infectious disease control", *International Journal of Health Geographics*, vol. 10, nr. 1/2011, p. 35, on-line la adresa https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3123167/ (accesată cel mai recent la 25 martie 2019).
15. Michelle V. Hauge, Mark D. Stevenson, D. Kim Rossmo și Steven C. Le Comber, „Tagging Banksy: using geographic profiling to investigate a modern art mystery", *Journal of Spatial Science*, vol. 61, nr. 1/2016, pp. 185–90, on-line la adresa http://www.tandfonline.com/doi/abs/10.1080/14498596.2016.1138246 (accesată cel mai recent la 25 martie 2019).
16. Raymond Dussault, „Jack Maple: betting on intelligence", *Government Technology*, 31 martie 1999, on-line la adresa

http://www.govtech.com/featured/Jack-Maple-Betting-on-Intelligence.html (accesată cel mai recent la 25 martie 2019).

17. *Ibidem.*
18. *Ibid.*
19. Nicole Gelinas, „How Bratton's NYPD saved the subway system", *New York Post*, 6 august 2016, on-line la adresa http://nypost.com/2016/08/06/how-brattons-nypd-saved-the-subway-system/ (accesată cel mai recent la 25 martie 2019).
20. Raymond Dussault, *op. cit.*
21. Andrew Guthrie Ferguson, „Predictive policing and reasonable suspicion", *Emory Law Journal*, vol. 62, nr. 2/2012, p. 259, on-line la adresa http://law.emory.edu/elj/content/volume-62/issue-2/articles/predicting-policing-and-reasonable-suspicion.html (accesată cel mai recent la 25 martie 2019).
22. Lawrence W. Sherman, Patrick R. Gartin și Michael E. Buerger, „Hot spots of predatory crime: routine activities and the criminology of place", *Criminology*, vol. 27, nr. 1/1989, pp. 27–56, on-line la adresa http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.1989.tb00862.x/abstract (accesată cel mai recent la 25 martie 2019).
23. Toby Davies și Shane D. Johnson, „Examining the relationship between road structure and burglary risk via quantitative network analysis", *Journal of Quantitative Criminology*, vol. 31, nr. 3/2015, pp. 481–507, on-line la adresa http://discovery.ucl.ac.uk/1456293/5/Johnson_art%253A10.1007%252Fs10940-014-9235-4.pdf (accesată cel mai recent la 25 martie 2019).
24. Michael J. Frith, Shane D. Johnson și Hannah M. Fry, „Role of the street network in burglars' spatial decision-making", *Criminology*, vol. 55, nr. 2/2017, pp. 344–76, on-line la adresa http://onlinelibrary.wiley.com/doi/10.1111/1745-9125.12133/full (accesată cel mai recent la 25 martie 2019).

25. Spencer Chainey, *Predictive Mapping (Predictive Policing)*, JDI Brief, Jill Dando Institute of Security and Crime Science, University College Londra, 2012, on-line la adresa http://discovery.ucl.ac.uk/1344080/3/JDIBriefs_PredictiveMappingSChaineyApril2012.pdf (accesată cel mai recent la 25 martie 2019).
26. *Ibidem.*
27. Un mic avertisment. Algoritmul PredPol în sine nu este accesibil publicului. Experimentul la care ne referim în carte a fost realizat de aceiași matematicieni care au creat PredPol, cu ajutorul unei tehnici care se potrivește cu modul în care este descris software-ul brevetat. Toate indiciile sugerează că amândoi algoritmii sunt la fel, dar, strict vorbind, nu putem fi *absolut* siguri.
28. G.O. Mohler, M.B. Short, Sean Malinowski, Mark Johnson, G.E. Tita, Andrea L. Bertozzi și P.J. Brantingham, „Randomized controlled field trials of predictive policing", *Journal of the American Statistical Association*, vol. 110, nr. 512/2015, pp. 1399–1411, on-line la adresa http://www.tandfonline.com/doi/abs/10.1080/01621459.2015.1077710 (accesată cel mai recent la 25 martie 2019).
29. Departamentul de analiză al serviciului administrativ al Poliției din Kent, *PredPol Operational Review*, 2014, on-line la adresa http://www.statewatch.org/docbin/uk-2014-kent-police-predpol-op-review.pdf (accesată cel mai recent la 25 martie 2019).
30. G.O. Mohler, M.B. Short, Sean Malinowski, Mark Johnson, G.E. Tita, Andrea L. Bertozzi și P.J. Brantingham, *op. cit.*
31. Departamentul de analiză al serviciului administrativ al Poliției din Kent, Marea Britanie, *PredPol Operational Review: Initial Findings*, 2013, on-line la adresa https://www.whatdotheyknow.com/request/181341/response/454199/attach/3/13%2010%20888%20Appendix.pdf (accesată cel mai recent la 25 martie 2019).
32. *Ibidem.*

33. Acesta nu era PredPol, ci un algoritm mai simplu care utiliza tot ideea efectelor de „semnal" și „intensificare". Vezi Matthew Fielding și Vincent Jones, „Disrupting the optimal forager: predictive risk mapping and domestic burglary reduction in Trafford, Greater Manchester", *International Journal of Police Science and Management*, vol. 14, nr. 1/2012, pp. 30–41.
34. Joe Newbold, '"Predictive policing", "preventative policing" or "intelligence led policing". What is the future?', Proiect consultativ prezentat spre evaluare de către programul de MBA Warwick, Warwick Business School, 2015.
35. Informații din 2016: COMPSTAT, *Citywide Profile 12/04/16–12/31/16*, on-line la adresa http://assets.lapdonline.org/assets/pdf/123116cityprof.pdf (accesată cel mai recent la 25 martie 2019).
36. Ronald V. Clarke și Mike Hough, *Crime and Police Effectiveness*, Home Office Research Study nr. 79, HMSO, Londra, 1984, on-line la adresa https://archive.org/stream/op1276605-1001/op1276605-1001_djvu.txt (accesată cel mai recent la 25 martie 2019), apud Tom Gash, *Criminal: The Truth about Why People Do Bad Things*, Allen Lane, Londra, 2016.
37. Departamentul de analiză al serviciului administrativ al Poliției din Kent, *op. cit.*
38. PredPol, „Recent examples of crime reduction", 2017, on-line la adresa http://www.predpol.com/results/ (accesată cel mai recent la 25 martie 2019).
39. Aaron Shapiro, „Reform predictive policing', *Nature*, vol. 541, nr. 7638/25 ianuarie 2017, on-line la adresa http://www.nature.com/news/reform-predictive-policing-1.21338 (accesată cel mai recent la 25 martie 2019).
40. Chicago Data Portal, Strategic Subject List, on-line la adresa https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List/4aki-r3np (accesată cel mai recent la 25 martie 2019).

41. Jessica Saunders, Priscilla Hunt şi John Hollywood, „Predictions put into practice: a quasi-experimental evaluation of Chicago's predictive policing pilot", *Journal of Experimental Criminology*, vol. 12, nr. 3/2016, pp. 347–371.
42. Copblock, „Innocent man arrested for robbery and assault, spends two months in Denver jail", 28 aprilie 2015, on-line la adresa https://www.copblock.org/122644/man-arrested-for-robbery-assault-he-didnt-commit-spends-two-months-in-denver-jail/ (accesată cel mai recent la 25 martie 2019).
43. *Ibidem.*
44. Ava Kofman, „How a facial recognition mismatch can ruin your life", *The Intercept*, 13 octombrie 2016.
45. *Ibidem.*
46. Copblock, „Denver police, "Don't fck with the biggest gang in Denver" before beating man wrongfully arrested – TWICE!!", 30 ianuarie 2016, on-line la adresa https://www.copblock.org/152823/denver-police-fck-up-again/ (accesată cel mai recent la 25 martie 2019).
47. De fapt, povestea lui Talley este chiar mai răscolitoare decât am prezentat-o aici în rezumat. După ce a petrecut două luni în închisoare, când a fost arestat iniţial, Talley a fost eliberat fără să i se aducă vreo acuzaţie. Un an mai târziu – în momentul în care el trăia deja într-un adăpost – a fost arestat pentru a doua oară. De această dată acuzaţiile nu au fost retrase, iar FBI a depus mărturie împotriva sa. Procesul a încetat atunci când funcţionara de la bancă şi-a dat seama că lui Talley îi lipseau aluniţele pe care le observase pe mâinile jefuitorului când acesta le întinsese peste ghişeu, şi a declarat în instanţă: „Acesta nu este individul care m-a jefuit." Acum Talley cere despăgubiri de zece milioane de dolari. Pentru relatarea completă, vezi Ava Kofman, *op. cit.*
48. Justin Huggler, „Facial recognition software to catch terrorists being tested at Berlin station", *Telegraph*, 2 august 2017, on-line la adresa http://www.telegraph.co.uk/

news/2017/08/02/facial-recognition-software-catch-terrorists-tested-berlin-station/ (accesată cel mai recent la 25 martie 2019).
49. David Kravets, „Driver's license facial recognition tech leads to 4,000 New York arrests", *Ars Technica*, 22 august 2017, on-line la adresa https://arstechnica.com/tech-policy/2017/08/biometrics-leads-to-thousands-of-a-ny-arrests-for-fraud-identity-theft/ (accesată cel mai recent la 25 martie 2019).
50. Ruth Mosalski, „The first arrest using facial recognition software has been made", *Wales Online*, 2 iunie 2017, on-line la adresa http://www.walesonline.co.uk/news/local-news/first-arrest-using-facial-recognition-13126934 (accesată cel mai recent la 25 martie 2019).
51. Sebastian Anthony, „UK police arrest man via automatic face recognition tech", *Ars Technica*, 6 iunie 2017, on-line la adresa https://arstechnica.com/tech-policy/2017/06/police-automatic-face-recognition (accesată cel mai recent la 25 martie 2019).
52. David White, Richard I. Kemp, Rob Jenkins, Michael Matheson și A. Mike Burton, „Passport officers' errors in face matching", *PLOSOne*, 18 august 2014, on-line la adresa http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0103510s6 (accesată cel mai recent la 25 martie 2019).
53. Teghan Lucas și Maciej Henneberg, „Are human faces unique? A metric approach to finding single individuals without duplicates in large samples", *Forensic Science International*, vol. 257/decembrie 2015, pp. 514e1–514.e6, on-line la adresa http://www.sciencedirect.com/science/article/pii/S0379073815003758 (accesată cel mai recent la 25 martie 2019).
54. Zaria Gorvett, „You are surprisingly likely to have a living doppelganger", BBC Future, 13 iulie 2016, on-line la adresa http://www.bbc.com/future/story/20160712-you-

are-surprisingly-likely-to-have-a-living-doppelganger (accesată cel mai recent la 25 martie 2019).
55. „Eyewitness misidentification", The Innocence Project, on-line la adresa https://www.innocenceproject.org/causes/eyewitness-misidentification (accesată cel mai recent la 25 martie 2019).
56. Douglas Starr, „Forensics gone wrong: when DNA snares the innocent", *Science*, 7 martie 2016, on-line la adresa http://www.sciencemag.org/news/2016/03/forensics-gone-wrong-when-dna-snares-innocent (accesată cel mai recent la 25 martie 2019).
57. Asta nu înseamnă că, dacă se apelează la ADN, falsele identificări sunt excluse – ele se mai întâmplă; înseamnă numai că ai o armă puternică de partea ta pentru a le face să fie cât mai rare cu putință.
58. Richard W. Vorder Bruegge, *Individualization of People from Images,* Divizia de Tehnologie Operațională a FBI, Unitatea de analiză criminalistică audio, video și imagine, Quantico, 12 decembrie 2016, on-line la adresa https://www.nist.gov/sites/default/files/documents/2016/12/12/vorderbruegge-face.pdf (accesată cel mai recent la 25 martie 2019).
59. Lance Ulanoff, „The iPhone X can't tell the difference between twins", *Mashable UK*, 31 octombrie 2017, on-line la adresa http://mashable.com/2017/10/31/putting-iphone-x-face-id-to-twin-test/A87kA26aAqqQ (accesată cel mai recent la 25 martie 2019).
60. Kif Leswing, „Apple says the iPhone X's facial recognition system isn't for kids", *Business Insider UK*, 27 septembrie 2017, on-line la adresa http://uk.businessinsider.com/apple-says-the-iphone-xs-face-id-is-less-accurate-on-kids-under-13-2017-9 (accesată cel mai recent la 25 martie 2019).
61. Andy Greenberg, „Watch a 10-year-old's face unlock his mom's iPhone X", *Wired*, 14 noiembrie 2017, on-line la adresa https://www.wired.com/

story/10-year-old-face-id-unlocks-mothers-iphone-x/ (accesată cel mai recent la 25 martie 2019).
62. „Bkav's new mask beats Face ID in "twin way": severity level raised, do not use Face ID in business transactions", Bkav Corporation, 27 noiembrie 2017, on-line la adresa http://www.bkav.com/dt/top-news/-/view_content/content/103968/bkav%EF%BF%BDs-new-mask-beats-face-id-in-twin-way-severity-level-raised-do-not-use-face-id-in-business-transactions (accesată cel mai recent la 25 martie 2019).
63. Mahmood Sharif, Sruti Bhagavatula, Lujo Bauer și Michael Reiter, „Accessorize to a crime: real and stealthy attacks on state-of-the-art face recognition", lucrare prezentată la Conferința ACM SIGSAC, 2016, on-line la adresa https://www.cs.cmu.edu/~sbhagava/papers/face-rec-ccs16.pdf (accesată cel mai recent la 25 martie 2019).
64. Ira Kemelmacher-Shlizerman, Steven M. Seitz, Daniel Miller și Evan Brossard, *The MegaFace Benchmark: 1 Million Faces for Recognition at Scale*, Computer Vision Foundation, 2015, on-line la adresa https://arxiv.org/abs/1512.00596 (accesată cel mai recent la 25 martie 2019).
65. „Half of all American adults are in a police face recognition database, new report finds", comunicat de presă, articol publicat de Facultatea de Drept a Universității din Georgetown, 18 octombrie 2016, https://www.law.georgetown.edu/news/press-releases/half-of-all-american-adults-are-in-a-police-face-recognition-database-new-report-finds.cfm (accesată cel mai recent la 25 martie 2019).
66. Josh Chin și Liza Lin, „China's all-seeing surveillance state is reading its citizens' faces", *Wall Street Journal*, 6 iunie 2017, on-line la adresa https://www.wsj.com/articles/the-all-seeing-surveillance-state-feared-in-the-west-is-a-reality-in-china-1498493020 (accesată cel mai recent la 25 martie 2019).

67. Daniel Miller, Evan Brossard, Steven M. Seitz și Ira Kemelmacher-Shlizerman, *The MegaFace Benchmark: 1 Million Faces for Recognition at Scale*, 2015, on-line la adresa https://arxiv.org/pdf/1505.02108.pdf (accesată cel mai recent la 25 martie 2019).
68. *Ibidem*.
69. MegaFace and MF2: Million-Scale Face Recognition, „Most recent public results", 12 martie 2017, http://megaface.cs.washington.edu/; „Leading facial recognition platform Tencent YouTu Lab smashes records in MegaFace facial recognition challenge", Cision PR Newswire, 14 aprilie 2017, on-line la adresa http://www.prnewswire.com/news-releases/leading-facial-recognition-platform-tencent-youtu-lab-smashes-records-in-megaface-facial-recognition-challenge-300439812.html (accesată cel mai recent la 25 martie 2019).
70. Dan Robson, „Facial recognition a system problem gamblers can't beat?", *TheStar.com*, 12 ianuarie 2011, https://www.thestar.com/news/gta/2011/01/12/facial_recognition_a_system_problem_gamblers_cant_beat.html (accesată cel mai recent la 25 martie 2019).
71. British Retail Consortium, *2016 Retail Crime Survey*, BRC, Londra, februarie 2017, on-line la adresa https://brc.org.uk/media/116348/10081-brc-retail-crime-survey-2016_all-graphics-latest.pdf (accesată cel mai recent la 25 martie 2019).
72. D&D Daily, *The D&D Daily's 2016 Retail Violent Death Report*, 9 martie 2017, on-line la adresa http://www.d-ddaily.com/archivesdaily/DailySpecialReport03-09-17F.htm (accesată cel mai recent la 25 martie 2019).
73. Joan Gurney, „Walmart's use of facial recognition tech to spot shoplifters raises privacy concerns", iQ Metrix, 9 noiembrie 2015, on-line la adresa http://www.iqmetrix.com/blog/walmarts-use-of-facial-recognition-tech-to-spot-shoplifters-raises-privacy-concerns (accesată cel mai recent la 25 martie 2019).

74. Andy Coghlan și James Randerson, „How far should fingerprints be trusted?", *New Scientist*, 14 sept. 2005, on-line la adresa https://www.newscientist.com/article/dn8011-how-far-should-fingerprints-be-trusted/ (accesată cel mai recent la 25 martie 2019).
75. Phil Locke, „Blood spatter – evidence?", *The Wrongful Convictions Blog*, 30 aprilie 2012, on-line la adresa https://wrongfulconvictionsblog.org/2012/04/30/blood-spatter-evidence/ (accesată cel mai recent la 25 martie 2019).
76. Michael Shermer, „Can we trust crime forensics?", *Scientific American*, 1 septembrie 2015, on-line la adresa https://www.scientificamerican.com/article/can-we-trust-crime-forensics/ (accesată cel mai recent la 25 martie 2019).
77. Consiliul Național de Cercetare al Academiei Naționale de Științe, *Strengthening Forensic Science in the United States: A Path Forward,* National Academies Press, Washington, DC, 2009, p. 7, on-line la adresa https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf (accesată cel mai recent la 25 martie 2019).
78. Colin Moynihan, „Hammer attacker sentenced to 22 years in prison", *New York Times*, 19 iulie 2017, on-line la adresa https://www.nytimes.com/2017/07/19/nyregion/hammer-attacker-sentenced-to-22-years-in-prison.html?mcubz=0 (accesată cel mai recent la 25 martie 2019).
79. Jeremy Tanner, „David Baril charged in hammer attacks after police-involved shooting", *Pix11*, 14 mai 2015, on-line la adresa http://pix11.com/2015/05/14/david-baril-charged-in-hammer-attacks-after-police-involved-shooting/ (accesată cel mai recent la 25 martie 2019).
80. „Long-time fugitive captured juggler was on the run for 14 years", informații FBI, 12 august 2014, on-line la adresa https://www.fbi.gov/news/stories/long-time-fugitive-neil-stammer-captured (accesată cel mai recent la 25 martie 2019).
81. Pei-Sze Cheng, „I-Team: use of facial recognition technology expands as some question whether rules are keeping

up", știre NBC prezentată în emisiunea *4NewYork* din 23 iunie 2015, on-line la adresa http://www.nbcnewyork.com/news/local/Facial-Recognition-NYPD-Technology-Video-Camera-Police-Arrest-Surveillance-309359581.html (accesată cel mai recent la 25 martie 2019).

82. Nate Berg, „Predicting crime, LAPD-style", *Guardian*, 25 iunie 2014, on-line la adresa https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report (accesată cel mai recent la 25 martie 2019).

CAPITOLUL 7 – *Arta*

1. Matthew J. Salganik, Peter Sheridan Dodds și Duncan J. Watts, „Experimental study of inequality and unpredictability in an artificial cultural market", *Science*, vol. 311/10 februarie 2006, p. 854, DOI [Digital Object Identifier]: 10.1126/science.1121066, on-line la adresa https://www.princeton.edu/~mjs3/salganik_dodds_watts06_full.pdf (accesată cel mai recent la 28 martie 2019).
2. Vezi site-ul dedicat MusicLab, on-line la adresa http://www.princeton.edu/~mjs3/musiclab.shtml (accesată cel mai recent la 28 martie 2019).
3. Kurt Kleiner, „Your taste in music is shaped by the crowd", *New Scientist*, 9 februarie 2006, on-line la adresa https://www.newscientist.com/article/dn8702-your-taste-in-music-is-shaped-by-the-crowd/ (accesată cel mai recent la 28 martie 2019).
4. Bjorn Carey, „The science of hit songs", *LiveScience*, 9 februarie 2006, on-line la adresa https://www.livescience.com/7016-science-hit-songs.html (accesată cel mai recent la 28 martie 2019).
5. „Vanilla, indeed", *True Music Facts Wednesday Blogspot*, 23 iulie 2014, on-line la adresa http://truemusicfactswednesday.blogspot.co.uk/2014/07/tmfw-46-vanilla-indeed.html (accesată cel mai recent la 28 martie 2019).

6. Matthew J. Salganik și Duncan J. Watts, „Leading the herd astray: an experimental study of self-fulfilling prophecies in an artificial cultural market", *Social Psychology Quarterly*, vol. 74, nr. 4/toamna 2008, p. 338, DOI: https://doi.org/10.1177/019027250807100404 (accesat cel mai recent la 28 martie 2019).
7. S. Sinha și S. Raghavendra, „Hollywood blockbusters and long-tailed distributions: an empirical study of the popularity of movies", *European Physical Journal B*, vol. 42, 2004, pp. 293–296, DOI: https://doi.org/10.1140/epjb/e2004-00382-7; http://econwpa.repec.org/eps/io/papers/0406/0406008.pdf (accesată cel mai recent la 28 martie 2019).
8. *„John Carter*: analysis of a so-called flop: a look at the box office and critical reaction to Disney's early tentpole release *John Carter*", *WhatCulture*, on-line la adresa http://whatculture.com/film/john-carter-analysis-of-a-so-called-flop (accesată cel mai recent la 28 martie 2019).
9. J. Valenti, „Motion pictures and their impact on society in the year 2000", discurs susținut la Midwest Research Institute, Kansas City, 25 aprilie 1978, p. 7.
10. William Goldman, *Adventures in the Screen Trade*, Warner, New York, 1983.
11. Sameet Sreenivasan, „Quantitative analysis of the evolution of novelty in cinema through crowdsourced keywords", *Scientific Reports* 3, articolul nr. 2758, 2013, actualizat la 29 ianuarie 2014, DOI: https://doi.org/10.1038/srep02758, https://www.nature.com/articles/srep02758 (accesat cel mai recent la 28 martie 2019).
12. Márton Mestyán, Taha Yasseri și János Kertész, „Early prediction of movie box office success based on Wikipedia activity big data", *PLoS ONE*, 21 august 2013, DOI: https://doi.org/10.1371/journal.pone.0071226 (accesat cel mai recent la 28 martie 2019).
13. Ramesh Sharda și Dursun Delen, „Predicting box-office success of motion pictures with neural networks", *Expert Systems with Applications*, vol. 30, nr. 2/2006, pp. 243–244,

DOI: https://doi.org/10.1016/j.eswa.2005.07.018 (accesat cel mai recent la 28 martie 2019); on-line la adresa https://www.sciencedirect.com/science/article/pii/S0957417405001399 (accesată cel mai recent la 28 martie 2019).
14. Banksy NY, „Banksy sells work for $60 in Central Park, New York – video", *Guardian*, 14 octombrie 2013, on-line la adresa https://www.theguardian.com/artanddesign/video/2013/oct/14/banksy-central-park-new-york-video (accesată cel mai recent la 28 martie 2019).
15. Bonhams, „Lot 12 Banksy: Kids on Guns", 2 iulie 2014, on-line la adresa http://www.bonhams.com/auctions/21829/lot/12/ (accesată cel mai recent la 28 martie 2019).
16. Charlie Brooker, „Supposing... subversive genius Banksy is actually rubbish', *Guardian*, 22 septembrie 2006, on-line la adresa https://www.theguardian.com/commentisfree/2006/sep/22/arts.visualarts (accesată cel mai recent la 28 martie 2019).
17. Gene Weingarten, „Pearls before breakfast: can one of the nation's greatest musicians cut through the fog of a DC rush hour? Let's find out", *Washington Post*, 8 aprilie 2007, on-line la adresa https://www.washingtonpost.com/lifestyle/magazine/pearls-before-breakfast-can-one-of-the-nations-great-musicians-cut-through-the-fog-of-a-dc-rush-hour-lets-find-out/2014/09/23/8a6d46da-4331-11e4-b47c-f5889e061e5f_story.html?utm_term=.a8c9b9922208 (accesată cel mai recent la 28 martie 2019).
18. Citatele provin dintr-o discuție particulară purtată cu Armand Leroi. Studiul la care se referă el este: Matthias Mauch, Robert M. MacCallum, Mark Levy și Armand M. Leroi, „The evolution of popular music: USA 1960–2010", *Royal Society Open Science*, 6 mai 2015, DOI: https://doi.org/10.1098/rsos.150081 (accesat cel mai recent la 28 martie 2019).
19. Citatele provin dintr-o discuție particulară purtată cu David Cope.

20. Acest citat a fost scurtat pentru mai multă concizie. Vezi Douglas Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid,* Penguin, Londra, 1979, p. 673.
21. George Johnson, „Undiscovered Bach? No, a computer wrote it", *New York Times*, 11 noembrie 1997.
22. Benjamin Griffin și Harriet Elinor Smith (editori), *Autobiography of Mark Twain*, vol. 3, Oakland, CA, și Londra, 2015, partea I, p. 103.
23. Leo Tolstoy, *What Is Art?*, Penguin, Londra, 1995, ediția princeps a fost publicată în 1897.
24. Douglas Hofstadter, *op. cit.*, p. 674.

Încheiere

1. Pentru povestea lui Rahinah Ibrahim, vezi https://www.propublica.org/article/fbi-checked-wrong-box-rahinah-ibrahim-terrorism-watch-list (link accesat cel mai recent la 28 martie 2019) și https://alumni.stanford.edu/get/page/magazine/article/?article_id=66231 (link accesat cel mai recent la 28 martie 2019).
2. GenPact, *Don't underestimate importance of process in coming world of AI*, 14 februarie 2018, on-line la adresa http://www.genpact.com/insight/blog/dont-underestimate-importance-of-process-in-coming-world-of-ai (accesată cel mai recent la 28 martie 2019).

MULȚUMIRI

Îmi imaginez că există și persoane cărora le este ușor să scrie. Cunoașteți genul – acelea care sar din pat înainte de ivirea zorilor și care până la prânz au scris deja un capitol și uită să coboare la cină pentru că sunt atât de absorbite de elanul lor creator, încât nu și-au dat seama că a trecut timpul.

Eu, una, cu siguranță nu sunt una dintre acestea.

Trecerea prin acest proces de creație a însemnat o luptă cotidiană cu acea parte din firea mea care nu vrea decât să stea pe canapea, ronțăind chipsuri și uitându-se la Netflix, și un război declarat cu valurile de panică și de îngrijorare pe care credeam că le-am lăsat în urmă în momentul în care mi-am terminat doctoratul. Nu pot să spun că *am scris* cartea aceasta, ci, mai degrabă, că am smuls-o din mine, lovind și țipând. Uneori, la propriu.

Prin urmare, sunt cu atât mai recunoscătoare remarcabilelor persoane care au fost dispuse să mă ajute pe parcurs. Extraordinarul colectiv de redacție, care a fost atât de generos cu timpul și ideile sale pe tot parcursul anului trecut: Susanna Wadeson, Quynh Do, Claire Conrad, Emma Parry, Gillian Somerscales, Emma Burton, Sophie Christopher, Hannah Bright, Caroline Saine și toți oamenii de la Janklow și Nesbit, Transworld și Norton care m-au ajutat din culise. Același lucru este valabil pentru Sue Rider, Kat Bee și Tom Copson: m-aș fi rătăcit de-a dreptul fără voi!

Mulțumiri din toată inima și celor pe care i-am intervievat, dintre care unii sunt citați în text, deși toți m-au ajutat să conturez

ideile din această carte: Jonathan Rowson, Nigel Harvey, Adam Benforado, Giles Newell, Richard Berk, Sheena Urwin, Steyve Colgan, Mandeep Dhami, Adrian Weller, Toby Davies, Rob Jenkins, Jon Kanevsky, Timandra Harkness, Dan Popple și echipa de poliție din West Midlands, Andy Beck, Jack Stilgoe, Caroline Rance, Paul Newman, Phyllis Illarmi, Armand Leoni, David Cope, Ed Finn, Kate Devlin, Shelia Hayman, Tom Chatwin, Carl Gombrich, Johnny Ryan, Jon Crowcroft și Frank Kelly.

Îi mai adaug aici pe Sue Webb și Debbie Enright de la Network Typing, precum și pe Sharon Richardson, Shruthi Rao și Will Storr, al căror ajutor a fost de nepreţuit pentru a reuși să dăm formă cărții de față. Pe lângă asta, în momentul în care am scris, în sfârșit, ceva care să semene cu niște fraze, James Fulker, Elisabeth Adlington, Brendan Maginnis, Ian Hunter, Omar Miranda, Adam Dennett, Michael Veale, Jocelyn Bailey, Cat Black, Tracy Fry, Adam Rutherford și Thomas Oléron Evans m-au ajutat să găsesc defectele cele mai evidente și să le îmblânzesc. Îi mulțumesc și lui Geoff Dahl, care, pe lângă sprijinul moral pe care mi l-a oferit pe parcursul întregului proces, a avut și o idee foarte bună pentru designul copertei.

Multe, multe mulțumiri comitetului meu de referenți: Elizabeth Cleverdon, Bethany Davies, Ben Dickson, Mike Downes, Charlie și Laura Galan, Katie Heath, Mia Kazi-Fornari, Fatah Ioualitene, Siobhan Mathers, Mabel Smaller, Ali Seyhun Saral, Jennifer Shelley, Edward Steele, Daniel Vesma, Jass Ubhi.

De asemenea, sunt nespus de recunoscătoare familiei mele pentru susținerea neabătută și statornica loialitate. Phil, Tracy, Natalie, Marge & Parge, Omar, Mike și Tania – ați avut mai multă răbdare cu mine decât am meritat adesea! (Nu o luați chiar *ad litteram*, totuși, pentru că s-ar putea să mai scriu o carte și voi avea din nou nevoie de ajutorul vostru, bine?)

Și, în cele din urmă, dar nu în ultimul rând, multe mulțumiri lui Edith! La drept vorbind, tu nu m-ai ajutat în niciun fel, dar nici nu aș fi vrut să fie altfel.

GLOSAR

A (se) autentifica (*log on/in* sau *sign in,* engl.) – „a deschide o sesiune de lucru cu un calculator aflat la distanță. În mod obișnuit, pentru a realiza această operație sunt necesare un nume de utilizator și o parolă" (Traian Anghel, *Dicționar de informatică,* Editura Corint, București, 2017). Operația inversă se numește deconectare (*log/ sign off,* engl.).

Adevăr întemeiat pe observație (*ground truth,* engl.) – informație empirică indubitabilă, obținută direct, prin observare, pe care se bazează un algoritm. Sintagma este folosită în numeroase domenii (tactică militară, geografie, cartografie, meteorologie etc.).

Agent software (*software agent,* engl.) – termen folosit în informatică pentru a desemna un program de computer capabil să acționeze (relativ) autonom și să manifeste următoarele caracterisitci: autonomie, pro-activitate, reactivitate, adaptabilitate, caracter cognitiv și social.

Algoritm *(algorithm,* engl.) – metodă (procedură) de calcul care descrie toate instrucțiunile sau toți pașii necesari care trebuie parcurși într-o anumită ordine pentru a obține un anumit rezultat.

Aplicație mobilă *(mobile application/mobile app,* engl.) – aplicație mobilă, program sau software menit să ruleze pe dispozitive mobile (telefon mobil, tabletă, smartwatch etc.).

Asistent (personal) virtual *(virtual assistant/intelligent personal assistant,* engl.) – asistent (personal) virtual, agent software, program care poate îndeplini diverse sarcini pe

baza unor comenzi vocale ale utilizatorului, putând chiar să răspundă folosind o voce sintetizată. Siri a companiei Apple, Alexa a Amazonului și Google Assistant sunt cei mai cunoscuți asistenți virtuali.

Bug (engl.) – „(deficiență, hibă), eroare de proiectare sau de programare care are drept consecință funcționarea incorectă a unui produs software" (Traian Anghel, *op. cit.*)

Cookie *(cookie HTTP, web cookie sau browser cookie, engl.; témoin de connexion,* fr.) – informație (codificată) trimisă de un site și stocată în computerul utilizatorului acestuia de către browserul său web, în timp ce utilizatorul accesează site-ul.

Broker de date sau de informații *(data broker,* engl.) – companie specializată în colectarea informațiilor despre persoane care există în surse publice sau private (de la recensăminte la registre ale adreselor, baze de date cu înregistrările de vehicule sau cu permisele de conducere, informații furnizate on-line de utilizatorii diverselor site-uri, registre ale tranzacțiilor bancare, istorice ale căutărilor on-line etc.), pe care le analizează și le însumează, pentru a realiza profiluri individuale (care pot conține sute și chiar mii de informații diverse despre o persoană sau un grup de persoane, de la gen, rasă, religie și până la venit lunar, probleme de sănătate etc.) pe care le vând apoi organizațiilor care doresc să organizeze campanii de reclame bine țintite *(targeted advertising).*

Data mapping (engl.) – descoperirea corespondențelor dintre informațiile deținute *(mise en correspondance de données,* fr.), termen folosit în informatică și în managementul informației pentru a defini crearea unor aplicații matematice menite să realizeze corespondențe între două sisteme de date (informații). Reprezintă primul pas în îndeplinirea unor sarcini de integrare de date diverse.

Data mining (engl.) – exploatarea bazelor de date *(exploration de données,* fr.), extragerea de cunoștințe din baze de

date, descoperirea de pattern-uri, de legături noi între diversele informații stocate.

Date (informatizate) (*data,* engl.) – informații sau „elemente de informație" (Traian Anghel, *op. cit.*).

Deep learning (engl.) – învățare profundă (numită în limba engleză și *deep structured learning sau hierachical learning),* modul în care învață o rețea neurală; ansamblu de metode de învățare automată specifice rețelelor neurale, capabile să creeze modele cu un înalt nivel de abstractizare din informațiile puse la dispoziție, folosite cu succes în domenii precum analizarea sunetelor, recunoaștere facială și vocală etc.

Deteriorarea relațiilor odată cu creșterea distanței (*distance decay,* engl.) – expresie folosită în geografie pentru a descrie modul în care distanța afectează interacțiunile (sociale, culturale etc.) dintre două centre, care postulează că interacțiunile dintre centre scad pe măsură ce distanța dintre ele crește. Efectul de deteriorare a relațiilor mai este cunoscut și sub numele de „prima lege a geografiei", formulată de geograful și cartograful american Waldo Tobbler, conform căreia: „toate lucrurile au relații cu toate celelalte, dar cele apropiate sunt mai legate unele de altele decât acelea aflate la distanță una de alta".

Digital Object Identifier (DOI) – șir de litere sau semne folosit pentru a identifica un document electronic. Nu se modifică, spre deosebire de URL-uri care se transformă de îndată ce se modifică structura serverului unde a fost publicat respectivul document on-line, așadar transmiterea DOI este o metodă de identificare mai bună decât prezentarea adresei URL.

Fals-pozitiv – (în medicină) rezultatul unui test care atestă prezența unui lucru inexistent; (în statistică) eroare de tip I (care neagă ipoteza nulă, care era corectă).

Feed (engl.) – „flux de știri sau canal de știri, care furnizează informații de sumar despre articolele noi publicate într-un site sau blog" (Traian Anghel, *op. cit.*)

Filtru de internet – numit și software de controlare a conținutului, reprezintă un ansamblu de tehnici care vizează limitarea accesului la anumite site-uri.

Informatică *(computer science,* engl.) – știința care studiază procesele de prelucrare automată a informațiilor (datelor), procese care pot fi, la rândul lor, reprezentate ca informații (sau programe informatice) sau „modalitățile practice pentru implementarea informațiilor în sistemele de calcul" (Traian Anghel, *op. cit.*). Presupune folosirea algoritmilor pentru prelucrarea, stocarea și comunicarea informațiilor digitale. Informatica teoretică studiază teoria limbajelor formale (automatica, teoria computațională și teoria complexității, teoria grafurilor etc.).

Ipoteza nulă – ipoteză presupusă a fi adevărată dacă nu este infirmată statistic de rezultate.

Învățare automată *(machine learning,* ML, engl.; *apprentissage automatique,* fr.), domeniu al informaticii și ramură a studiilor dedicate inteligenței artificiale (IA) care cercetează algoritmii și modelele statistice care permit sistemelor informatice să îndeplinească o sarcină în lipsa unor instrucțiuni specifice, doar pe baza decelării unor modele și a realizării unor inferențe logice.

Listă de corespondență *(mailing list,* engl.) – listă care cuprinde adresele sau e-mailurile unor clienți potențiali cărora, în cadrul unor strategii de marketing direct, li se trimit produse sau informații despre acestea (prin intermediul newsletterelor, broșurilor diverse sau reclamelor).

Mass-marketing (engl.) – marketing de/în masă, care se adresează tuturor, fără a face nicio diferențiere între diversele segmente de piață.

Mașină *(machine,* engl.) – „termen echivalent cu calculator, gazdă sau sistem" (Traian Anghel, *op. cit.*).

Mesagerie instantanee (MI) (desemnată și prin anglicismul chat, pronunțat *cet) (instant messaging,* IM, engl.) – tehnologie care oferă transmitere instantanee de mesaj text pe internet. Skype sau Facebook Messenger sunt MI.

Micro-targetare *(micro-targeting,* engl.) – strategie folosită în general de partidele politice, pentru a trimite reclame create special pentru segmente foarte mici de alegători, care au fost grupați cu ajutorul unor procese de *data mining.*

Nișă de piață *(market niche,* engl.) – „segment de piață de dimensiuni reduse, sector foarte atractiv în cadrul unei piețe, care oferă oportunități pentru firmă (avantaje competitive și concurențiale) în ceea ce privește vânzările, rentabilitatea, lipsa concurenților puternici etc." (apud șef lucrări dr. Florin D. Frone, *Dicționar de marketing, f.a.).*

Prejudecată *(bias,* engl.) – predispoziția de a face unele alegeri care nu sunt fondate pe raționament; în cazul algoritmilor, acest termen denotă o eroare sistematică ce a fost introdusă prin intermediul unui eșantion care favorizează anumite rezultate. Spre exemplu, un algoritm de detectare a fețelor în poze ar putea avea bias către oamenii albi, întrucât eșantionul pe care a fost conceput nu conținea exemple cu oameni de culoare.

Provider de internet *(Internet Service Provider,* ISP, engl.) – numit și furnizor de servicii internet, reprezintă o companie care le oferă clienților săi conexiune și acces la internet.

Reclame țintite *(targeted advertising,* engl.) – formă de advertising on-line care folosește metode sofisticate pentru a putea să se adreseze doar persoanelor celor mai interesate de serviciul sau produsul căruia i se face reclamă. Selectarea potențialilor consumatori se face pe baza unor criterii diverse (gen, rasă, vârstă, statut economic, nivel de școlarizare, comportament la shopping, istoric al cumpărăturilor on-line etc.). Reprezintă opusul publicității nediferențiate, care este o strategie de *mass-marketing.*

Rețea neurală (artificială) *(artificial neural network, ANN,* engl.) – rețea compusă din neuroni artificiali, structură menită să proceseze informațiile a cărei sursă de inspirație au constituit-o rețelele neurale biologice și modul acestora de a structura informațiile. Sunt capabile să învețe din experiență.

Secvențierea ADN-ului *(sequencing DNA,* engl.) – procedură prin care se analizează codul ADN.

Știința datelor *(data science,* engl.) – domeniu de studiu multidisciplinar care folosește diverse metode științifice pentru extragerea informațiilor din baze de date bine structurate sau nestructurate. Concepte similare: *data mining* sau *big data.*

Thread (engl.) – fir de execuție, „în programare, flux de instrucțiuni care se execută în interiorul unui proces [...] care poate fi format din mai multe asemenea fire, care se execută în paralel, având însă în comun toate resursele principale caracteristice procesului." (Traian Anghel, *Dicționar de informatică);* (pe un forum) fir de discuție, subiect comentat de mai mulți utilizatori.

Urme digitale/informatice *(digital shadow sau digital footprint,* engl.) – totalitatea informațiilor despre o persoană decelabile din activitățile pe care aceasta le desfășoară în mediul on-line sau pe aplicațiile mobile (telefonie, servicii de mesagerie etc.): acțiuni diverse, produse achiziționate on-line, mesaje comunicate, conținuturi apreciate etc. Deși pot să nu prezinte un interes prea mare dacă rămân izolate, devin o sursă inestimabilă de informații atunci când sunt cumulate și interpretate, căci contribuie la crearea profilului psihologic al utilizatorului.

INDEX

23andMe 145–147
52Metro 227

Academia Națională de Științe a SUA 221
Acxiom 52
Adele 246
AF447 (zbor) 173, 175
Afigbo, Chukwuemeka 16
AI Music 246
Air France 173–176, 179
Alhambra 204
Alton Towers 38, 39
ALVINN 156–157
Alzheimer (maladie) 124–126, 145, 147
Amazon (motor de recomandări) 24, 229
Ancestry.com 147
Apple 73, 215
Audi 178, 183

Bainbridge, Lisanne 175–177, 180
Banksy (artist) 192, 237
Baril, David 222, 223
Bartlett, Jamie 69
Barwell, Clive 190–192
Bayes, teorema lui 159–164, 167
Bayes, Thomas 162
Beck, Andy 114, 129, 130
Bell, Joshua 238
Berk, Richard 90, 93, 99
Bonin, Pierre-Cédric 173–175
Bratton, Bill 194–195, 198
Brixton 75–77
Brooks, Christopher Drew 92, 93, 108
Brown, Joshua 177
Burgess, Ernest W. 82–83

Cambridge Analytica 62–67, 68, 72
CAMELYON16, 121, 122
Centaur Chess 257
China 112, 113, 217
Citroen DS19 154
Citymapper 41
Clubcard (Tesco) 45–46
Colgan, Steyve 202
Commodore 64 11

CompStat 195
Cope, David 242–244, 247
Corelogic 52
Creemers, Rogier 72
cutremure 198–199

Daimler 153, 171
DARPA 150–152, 155, 157, 158, 167
Davies, Toby 204, 205
Deep Blue 19–21, 23, 132, 257
DeepMind 139, 140
Dewes, Andreas 59
Dhami, Mandeep 106, 107, 108
Dodds, Peter 226
Douglas, Neil 211, 212
Dubois, căpitanul 179
Duggan, Mark 75
Dunn, Edwina 45

eBureau 52
Eckert, Svea 59, 60
Eppink, Richard 34–35
Epstein, Robert 31–32, 68
Equivant (fosta Northpointe) 99, 272
Experimente de Inteligență Muzicală (EMI) 242–245, 247–248

Face ID (Apple) 215
Facebook 16, 24, 25, 43, 58, 63–64, 67–68, 73
FaceFirst 221
FaceNet (Google) 217–218
FBI 191, 208, 209, 212, 214, 217, 222, 252
Financial Times 153
Firebird II 154
Firefox 73
Ford 66, 153, 171
Fuchs, Thomas 135
Fundația Clinton 66

Galton, Francis 143–145
GDPR (General Data Protection Regulation), 72
General Motors 154
Germania 59, 138
Goldman, William 232, 235
Google 31, 32, 36, 47, 58, 60, 100, 139, 140, 141, 170, 217, 218
Google Brain 130
Google Images 99
Google Maps 158
Google Search 23
Gogle Street-View 209
Google Translate 61
GPS 17, 29, 33, 151, 158, 159, 163, 176
Guerry, André-Michel 187–188

Hamm, John 134
Hammond, Philip 153
Harkness, Timandra 141
Hinton, Geoffrey 119
Hipocrate 111, 112
Hofstadter, Douglas 242, 248

Hugo, Christoph von 164, 165
Humby, Clive 45, 46, 47
Hume, David 236
HunchLab 205, 206, 223
Huntington (boală) 147

IA (inteligență artificială) 26, 28, 86, 118, 119, 239, 254
IBM 19, 20, 131, 132, 134, 135, 139, 256
Ibrahim, Rahinah 251–252
Innocence Project 213
Instagram 58, 222
Internet Movie Database (IMDb) 232

Jemaah Islam 252
Jemaah Islamiyah 252
Jennings, Ken 132
Jeopardy (concurs TV) 132–133, 135, 256
John Carter (film) 231
Johnson, Richard 76–78
Jones Beach 15
Jones, Robert 29–30, 42
Jukebox 246

Kadoodle 31–32
Kahneman, Daniel 102–103
Kanevsky, dr. Jonathan 127, 129
Kant, Immanuel 237
Kasparov, Garry 19–22, 30, 42, 132, 257
Kelly, Frank 119
Kerner, Winifred 241
Kernighan, Brian 12
Killingbeck 190–191

Larson, Steve 241
Legea lui Weber 105–106
Leibniz, Gottfried 236
Leroi, Armand 238–239
Li Yingyun 71
Lickel, Charles 132
LiDAR (Light Detection and Ranging) 157–159
Lista cu subiecți strategici 206, 255
„Lockdown" (52Metro) 227
London Bridge 222
London School of Economics (LSE) 169
Loomis, Eric 99
Lucas, Teghan 210

Maple, Jack 192–196
Marea Britanie 201–202, 205
Marx, Gary 224
McGrayne, Sharon Bertsch 161
Meehl, Paul 40
Mercedes 164–166
Millgarth 190–191
Mills, Tamara 136, 138
MIT Technology Review 135
Moses, Robert 15
Music Lab 227
Musk, Elon 180
MyHeritage 147

National Geographic 147
Navlab 154
Netflix 239–240
New England Journal of Medicine 128
New York City 233
New York Times 50, 153, 228, 241
Newman, Paul 167–169, 171

OKCupid 24
Ontario 219
Operațiunea Lynx 187, 190–191
Ordinul Surorilor de Notre Dame 124
Oxbotica 167

Palantir Technologies 52
PathAI 114, 129
Petrov, Stanislav 37–39
Pomerleau, Dean 156–157
Pratt, Gill 180
PredPol (PREDictive POLicing) 199–201, 203–206, 223
programul Crimewatch 186
Proiectul Eureka Prometheus 155
ProPublica 94

radar 157–159
RAND Corporation 206
război nuclear 37
Războiul Rece 37
Rhode Island 89
Robinson, Nicholas 75–78
Rossmo, Kim 186–187
Rotten Tomatoes (website) 229, 232
Royal Free NHS Trust 139
Rutherford, Adam 147

Safari (browser) 73

Sainsbury's 46
Salganik, Matthew 226, 228
Schmidt, Eric 47
Science (revistă) 32, 166
Scunthorpe 16
Semmelweis, Ignaz 112
Sesame Credit 70, 218
Slash X (bar) 150, 152
Snowdon, David 124
Sorensen, Alan 229
Spotify 226, 229, 239–240
Spotify Discover 240
Sreenivasan, Sameet 232–233
Stammer, Neil 222
Stilgoe, Jack 169–171
SUA 252

Talley, Steve 207
Tesco 45–47
Tesla 177, 181–182
Testul Magiei 254
Thiel, Peter 52
Timberlake, Justin 225
Timberlake, Justin (artist) 226–227
Tolstoi, Lev 247

TomTom 23, 30
Toyota 180, 183
Twain, Mark 247
Twitter 25, 52, 58, 60, 64

Uber 25, 178
Uniunea Libertăților Civile Americane (ACLU) 34
Uniunea Sovietică 224
Universitatea din Texas, Centrul de Cancer M. D. Anderson 134
URL-uri 60–61

Valenti, Jack 232
Vanilla (formație) 229
Volvo 168, 183
Volvo XC90 183

Walmart 221
Walt Disney 232
Warhol, Andy 237
Washington Post 238
Waterhouse, Heidi 57
Watson (calculator IBM) 131–136, 138, 143, 256
Watts, Duncan 226
Waymo 170–171
Waze 41–42
Williams, Pharrell 246
Windows XP 141
Wired 53, 177

Xing.com 60

Zaghba, Youssef 222–223
Zilly, Paul 91
Zuckerberg, Mark 43
ZX Spectrum 11

În curând în colecția

CORINT FUTURE

Jamie Susskind

POLITICA VIITORULUI

Cum să trăim împreună într-o lume transformată de tehnologie

Traducere din limba engleză
Adina Ihora

— Fragment —

INTRODUCERE

Viitorul ne urmărește. Este tot timpul în așteptare, abia ascuns vederii, pândind de după colț sau după următorul răsărit. Nu știm sigur ce formă va lua. Adesea ne surprinde complet nepregătiți.

În prezent, mulți dintre noi au impresia că se apropie o perioadă de mari frământări. Lumea pare să se schimbe mai repede decât reușim să o înțelegem. Ne străduim adesea să ne explicăm evenimente politice care, cu numai câțiva ani în urmă, ar fi fost de neînchipuit. Uneori, nu dispunem nici măcar de cuvintele potrivite care să le descrie. În sinea noastră știm, însă, că acesta nu este decât începutul.

Premisa de la care pornește cartea de față este că progresele neîncetate din știință și tehnologie urmează să transforme felul în care trăim, având consecințe profunde și totodată înspăimântătoare în ceea ce privește politica. Nu suntem încă pregătiți – intelectual, filosofic sau moral – pentru lumea pe care o creăm. În următorii câțiva zeci de ani, vechile moduri de gândire, care ne-au fost utile secole sau milenii de-a rândul, vor fi puse sub semnul întrebării. Noi dezbateri, controverse, mișcări și ideologii vor ajunge în prim-plan. Unele dintre convingerile noastre cele mai profunde vor fi revizuite sau cu totul abandonate. Va trebui să ne imaginăm împreună, din nou, ce înseamnă să fim liberi sau egali, ce înseamnă să deținem putere sau o proprietate, chiar și ce înseamnă ca un sistem politic să fie democratic. Politica din viitor nu va semăna aproape deloc cu politica din trecut.

Politica secolului al XX-lea a fost dominată de o întrebare esențială: cât de mult din viața noastră colectivă ar trebui să fie stabilită de stat și cât din aceasta ar trebui lăsată în seama pieței și a societății civile? Pentru generația care se apropie în prezent de maturitatea politică, discuția va fi diferită: în ce măsură ar trebui ca viețile noastre să fie dirijate și controlate de puternice sisteme digitale – și în ce condiții? Această întrebare se află în centrul cărții *Politica viitorului*.

În următoarele câteva decenii se preconizează că vom dezvolta sisteme informatice uimitoare, dintre care unele vor rivaliza cu și, poate, vor depăși ființele umane într-o paletă largă de atribuții, chiar și fără să ajungă la o „inteligență" ca a noastră. Aceste sisteme vor fi integrate în lumea fizică, ascunse în structuri și obiecte pe care nu am fost obișnuiți să le considerăm drept tehnologie. Din ce în ce mai multe informații despre ființele umane – ce facem, unde mergem, ce gândim, ce spunem, ce simțim – vor fi colectate și înregistrate ca date, apoi vor fi sortate, stocate și procesate digital. Pe termen lung, diferențele dintre om și mașină, on-line și off-line, virtual și real, se vor estompa.

Această transformare va aduce civilizației câteva mari beneficii. Viețile noastre vor fi îmbogățite de noi modalități de a ne juca, de a lucra, a călători, a face cumpărături, a învăța, a crea, a ne exprima, a păstra legătura cu ceilalți, a întâlni persoane necunoscute, a coordona acțiuni, a ne menține în formă și a găsi înțelesuri. Pe termen lung, s-ar putea să reușim să ne perfecționăm mintea și trupul, făcându-le de nerecunoscut și eliberându-ne de limitările biologiei umane.

În același timp, însă, unele tehnologii vor ajunge să dețină o mare putere asupra noastră. Vor putea să ne forțeze să ne comportăm într-un anumit fel, cum ar fi (ca să luăm un exemplu simplu) vehiculele autonome, care refuză efectiv să depășească limita de viteză prestabilită. Alte tehnologii vor fi puternice datorită informațiilor pe care le colectează despre noi. Simplul fapt că ne știm supravegheați ne face mai

puțin predispuși să înfăptuim lucruri percepute ca rușinoase, păcătoase sau greșite. Alte tehnologii vor filtra lumea pe care o vedem, stabilind ce să cunoaștem, modelând felul în care gândim, influențând ceea ce simțim și determinând astfel modul în care acționăm.

Cei care controlează aceste tehnologii incredibile ne vor controla din ce în ce mai mult și pe noi. Ei vor deține *putere*, ceea ce înseamnă că vor dispune de capacitatea constantă și extinsă de a ne determina să facem lucruri importante, pe care, altfel, nu le-am face. Din ce în ce mai mult, ei ne vor limita *libertatea*, stabilind ce este permis și ce este interzis. Vor stabili viitorul *democrației*, ducând la dezvoltarea sau la degradarea acesteia. Iar algoritmii lor vor lua decizii în probleme vitale de *justiție socială*, alocând beneficii și împărțindu-ne pe ierarhii, în funcție de statutul nostru și de respectul pe care îl inspirăm.

În consecință, autoritățile politice – statele, în general – vor dispune de mai multe instrumente de control decât oricând, iar marile companii tehnologice se vor bucura, la rândul lor, de putere la un nivel care va eclipsa orice altă entitate economică din timpurile moderne. Pentru a face față acestor noi provocări, vom avea nevoie de modernizarea radicală a ideilor noastre politice. [...]

URMĂTOAREA MARE SCHIMBARE

Trăim deja într-o perioadă de profundă neliniște politică. În fiecare zi aflăm știri despre războaie civile sângeroase, deplasări masive ale populației, naționalism etnic, violență sectară, extremism religios, schimbări climatice, tulburări economice, globalizare derutantă, inegalitate crescândă și o diversitate de alte evenimente problematice, prea urâte ca să le amintim aici. S-ar părea că lumea nu o duce prea bine și că discursurile noastre publice nu sunt la înălțimea așteptărilor. Se manifestă o amplă neîncredere și un mare dispreț față

de elitele politice. Două recente exerciții democratice de masă din lumea vorbitoare de limbă engleză, alegerile prezidențiale din SUA din 2016 și referendumul privind Brexit din Marea Britanie, au fost pline de resentimente, chiar și după standardele de nefericire obișnuite, facțiunile opuse străduindu-se nu numai să-și învingă rivalii, ci și să-i nimicească. Ambele au fost câștigate de partea care promitea să distrugă vechea ordine. Niciunul dintre aceste exerciții nu a adus după sine împăcare sau mulțumire. Din ce în ce mai mult, după cum remarca Barack Obama la sfârșitul mandatului său prezidențial, „totul este adevărat și nimic nu este adevărat". Devine din ce în ce mai dificil pentru cetățenii obișnuiți (indiferent de apartenența lor politică) să facă distincția dintre adevăr și impostură, realitate și zvon, semnal și zgomot. Mulți chiar au renunțat să mai încerce. Suntem tentați să rămânem pe poziții și să ținem piept furtunii din prezent, fără să ne gândim prea mult la viitor.

Însă aceasta ar fi o greșeală.

Dacă cele mai importante previziuni despre viitorul tehnologiei se apropie de adevăr, atunci schimbarea care se întrevede ar putea fi cel puțin tot atât de importantă pentru umanitate precum revoluția industrială, revoluția agrară sau chiar inventarea limbii vorbite. Prin comparație, multe dintre problemele actuale vor fi eclipsate. Gândiți-vă la efectul pe care l-a avut deja tehnologia în viața noastră – asupra felului în care lucrăm, comunicăm, ne tratăm afecțiunile, facem exerciții fizice, mâncăm, învățăm și socializăm –, iar apoi amintiți-vă că, din perspectivă istorică, era digitală are o vârstă de numai câteva secunde. În total, 99,5% din existența umană s-a petrecut în epoca paleolitică, apărută cu aproximativ trei milioane de ani în urmă, când oamenii au început să folosească unelte primitive. Epoca aceea a luat sfârșit în urmă cu vreo 12 000 de ani, odată cu ultima glaciațiune. În timpul acestei lungi perioade crepusculare, oamenii nu au remarcat aproape niciun fel de schimbare culturală. „Lumea în care oamenii intrau la

naştere era aceeaşi cu cea pe care o părăseau în momentul morţii." Dacă vă gândiţi că cele mai vechi civilizaţii umane au apărut în urmă cu vreo 5 000 de ani, atunci cei şaptezeci şi ceva de ani pe care i-am trăit alături de calculatoarele moderne, cei treizeci şi ceva de ani de când dispunem de World Wide Web şi cei zece ani de când utilizăm telefoanele inteligente nu par nicidecum nişte perioade îndelungate. Pe când timpul se scurge în mod liniar, numeroasele progrese din tehnologia digitală au loc exponenţial, viteza schimbării fiind, de la un an la altul, tot mai mare.

Nu avem niciun indiciu cu privire la viitor, aşadar încercarea de a-l anticipa este inerent riscantă şi dificilă. Îi admir pe cei care se străduiesc să facă acest lucru în mod riguros şi m-am inspirat pe larg din activitatea lor la întocmirea acestei cărţi. Dar ca să fim realişti, ar trebui să începem prin a recunoaşte că astfel de predicţii se înşală adesea foarte tare. Probabil că o mare parte a viitorului anticipat în paginile următoare nu va deveni niciodată realitate şi alte evenimente, complet neprevăzute, îşi vor face în schimb apariţia, luându-ne prin surprindere. Acestea fiind spuse, cred că se pot face supoziţii de bun-simţ, în cunoştinţă de cauză, despre felul în care ar putea arăta viitorul în baza a ceea ce cunoaştem astăzi despre tendinţele actuale din ştiinţă, tehnologie şi politică. Riscul cel mai mare ar fi să nu încercăm în niciun fel să anticipăm viitorul.

Se povesteşte despre o întâlnire dintre omul de stat din epoca victoriană William Gladstone şi sclipitorul om de ştiinţă Michael Faraday. Acesta din urmă încerca să-i explice lui Gladstone activitatea lui revoluţionară în domeniul electricităţii, dar Gladstone nu părea impresionat.

— Dar la ce este bună chestia asta? a întrebat cel din urmă, cu o frustrare crescândă. La ce *foloseşte* ea?

— Ei bine, domnule, i-a răspuns Faraday, ajuns la limita răbdării, sunt şanse mari ca, în curând, să puteţi pune un impozit pe asta.

Multor inovatori, asemenea lui Faraday, le este greu să explice implicațiile sociale și practice ale activității lor. Iar noi, ceilalți, asemenea lui Gladstone, respingem mult prea adesea tehnologiile pe care încă nu le înțelegem. Poate fi greu de observat importanța politică a invențiilor care, la prima vedere, par să nu aibă nimic în comun cu politica. Când ne aflăm în fața unui *gadget* sau a unei aplicații noi, nu ne gândim mai întâi la implicațiile politice pe care le vor avea acestea. Vrem, în schimb, să știm: la ce folosește asta? Cât costă? De unde îmi pot procura și eu așa ceva? Nu este deloc surprinzător. În general, ne întâlnim cu tehnologia în calitate de consumatori, însă această atitudine relativ obtuză trebuie să se schimbe. Trebuie să dedicăm noilor tehnologii ale puterii aceeași examinare atentă și același scepticism pe care le-am rezervat întotdeauna politicienilor puternici. Tehnologia nu ne afectează numai în calitate de consumatori, ci și de cetățeni. În secolul al XXI-lea, *digitalul este politizat.*

Această carte li se adresează, pe de o parte, acelor Gladstone care vor să înțeleagă mai mult din tehnologie, iar pe de altă parte, acelor Faraday care vor să înțeleagă mai bine importanța politică a activității lor. Însă cea mai mare parte a cărții se adresează oamenilor obișnuiți, care doresc să priceapă viitorul ceva mai bine, dacă nu din alte motive, măcar pentru a-i responsabiliza pe toți acei Gladstoni și Faraday.

INGINERII FILOSOFI

Reflectați puțin la fragmentul de mai jos:

> Cinste celor nebuni. Celor inadaptați. Celor rebeli. Celor zurbagii. Cuielor rotunde din găurile pătrate. Celor care văd lucrurile altfel. Celor cărora nu le plac regulile. Și care nu au niciun respect față de starea actuală a lucrurilor. Poți să-i citezi, să nu fii de acord cu ei, să-i lauzi sau să-i denigrezi. Singurul lucru pe care nu-l poți face este să-i ignori. Pentru

că ei schimbă lucrurile. Fac rasa umană să avanseze. Deși unii i-ar putea considera nebuni, noi vedem în ei niște genii. Pentru că oamenii care sunt destul de nebuni să creadă că pot să schimbe lumea sunt cei care chiar o schimbă.

Acestea nu sunt cuvintele unui politician. Sunt extrase din postsincronul de la „Think Different"*, o reclamă a companiei Apple din 1997, în care erau incluse imagini ale unor rebeli celebri, printre care Mahatma Gandhi și Martin Luther King. Reclama prezintă o anumită viziune asupra lumii, larg răspândită printre antreprenorii din domeniul tehnologiei, și anume că activitatea lor are și o importanță filosofică, pe lângă cea comercială. „În Silicon Valley este ceva obișnuit", arată Jaron Lanier**, „ca persoanele foarte tinere, cu un start-up amenajat în vreun garaj, să declare că țelul lor este să schimbe din temelii cultura umană, la scală mondială, în numai câțiva ani, și că nu sunt încă pregătiți să-și facă griji în privința banilor pentru că acumularea unei mari averi este o chestiune măruntă, care se va rezolva de la sine." Un astfel de mod de gândire are ceva atrăgător, în parte pentru că sugerează că aceste companii tehnologice ar putea să nu fie atât de rapace pe cât sunt considerate îndeobște. Iar premisa de la care se pornește este una corectă: tehnologiile digitale chiar *au* o capacitate uimitoare de a schimba lumea. Comparați afirmațiile următoare:

„Filosofii nu au făcut decât să *interpreteze* lumea în diferite moduri; ideea este, însă, *schimbarea* sa."

„Noi nu *analizăm* o lume, ci o *făurim*."

Prima datează din 1845 și îi aparține lui Karl Marx – *Teze despre Feuerbach*. A servit drept slogan revoluționarilor politici mai mult de o sută de ani de la publicarea sa. Cea de-a doua îi aparține manieratului Tim Berners-Lee, inventatorul

* Gândește diferit (n. tr.)

** Jaron Lanier (n. 3 mai 1960, New York, SUA), programator, muzician și părintele realității virtuale (n. tr.).

World Wide Web. Marx și Berners-Lee nici că ar putea fi mai diferiți în ceea ce privește gândirea politică, temperamentul sau stilul lor de păr facial. Însă ceea ce au în comun – în plus față de faptul că au schimbat mersul istoriei omenirii – este convingerea că există o deosebire între *a aduce* o schimbare și doar a te *gândi* la schimbare sau a *studia* schimbarea. Din această perspectivă, departe de a fi o prezență spectrală asupra căreia nu avem niciun control, viitorul este ceva pe care îl elaborăm și îl construim.

„Noi nu suntem filosofi experimentaliști", spunea cândva Berners-Lee, „suntem ingineri filosofi." Este un mod practic și participativ de a privi viața; unul mai familiar constructorilor și inventatorilor, decât cadrelor universitare îmbrăcate în stofă tweed sau filosofilor cu pulovere pe gât. Se întâmplă totodată să fie și mentalitatea definitorie a epocii în care trăim. În prezent, cele mai importante revoluții nu au loc în departamentele de filosofie, nici chiar în parlamente sau în agora orașelor, ci în laboratoare, centre de cercetare, companii tehnologice și centre de date. Majoritatea presupun dezvoltări ale tehnologiilor digitale. Și, totuși, aceste avansuri extraordinare se petrec într-un alarmant climat de izolare culturală și intelectuală. Cu puține excepții, între arte și știință se cască o prăpastie. Filosofia politică și politica socială apar rareori în programele de studii universitare din domeniul științei, tehnologiei, ingineriei și matematicii. Iar dacă îl întrebi pe studentul clasic la arte liberale cum funcționează un calculator, este puțin probabil să obții un răspuns elaborat.

În companiile tehnologice, extrem de puțini ingineri sunt însărcinați să se gândească intens la consecințele sistemice ale activității lor. Celor mai mulți li se dau spre rezolvare probleme tehnice concrete. În sectorul tehnologic, inovația este, până la urmă, motivată de obținerea de profit, chiar dacă investitorii sunt pregătiți să adopte o abordare de genul „o idee bună la început, iar profitul mai târziu". Nu este o observație critică: înseamnă doar că nu există niciun motiv

pentru care a face bani și a face lumea mai bună să însemne întotdeauna unul și același lucru. De fapt, așa cum o arată multe dintre exemplele din această carte, există o sumedenie de dovezi care sugerează că tehnologia digitală este mult prea adesea concepută din perspectiva celor puternici și privilegiați.

Odată cu trecerea timpului, vom avea nevoie de mai mulți ingineri filosofi demni de acest nume. Și va deveni încă și mai important pentru noi, ceilalți, să interpretăm în mod critic activitatea companiilor de tehnologie, nu în ultimul rând deoarece cultura organizațională din domeniul tehnologiei este recunoscută pentru lipsa ei de diversitate. Aproximativ nouă din zece directori executivi din Silicon Valley sunt bărbați. În ciuda faptului că afro-americanii alcătuiesc cam 10% din absolvenții din IT și 14% din forța de muncă totală, aceștia ocupă sub 3% din posturile din domeniul informaticii în Silicon Valley. Și numeroase persoane din comunitatea tehnologică nutresc vederi politice ferme, care se situează mult în afara curentului general. Peste 44% din cei care au adoptat moneda Bitcoin în anul 2013, de exemplu, s-au declarat a fi „libertarieni sau anarho-capitaliști care se pronunță în favoarea eliminării statului".

Așa cum voi căuta să arăt în continuare, riscăm foarte mult atunci când delegăm chestiuni de importanță politică grupului infim de persoane care se întâmplă să fie însărcinate cu dezvoltarea tehnologiilor digitale la un moment dat. Acest lucru este adevărat indiferent dacă îi admiri pe inginerii filosofi din Silicon Valley sau dacă ești de părere că cei mai mulți „frați întru tehnologie" au rafinamentul politic al unui tranzistor. Avem nevoie de un cadru intelectual care să ne ajute să ne gândim limpede și cu simț critic la consecințele politice ale inovării digitale. Cartea aceasta speră să contribuie la un astfel de cadru, utilizând ideile și metodele teoriilor politice.

Scopul filosofiei, considera Isaiah Berlin, este întotdeauna același: să-i ajute pe oameni „să se înțeleagă pe ei înșiși și astfel să acționeze deschis, nu la întâmplare, nu în necunoștință de cauză." Acesta este și scopul nostru. Teoria politică își propune să înțeleagă politica prin intermediul conceptelor pe care le utilizăm pentru a ne referi la aceasta.Ce este *puterea*? Când ar trebui restricționată *libertatea* și pe ce considerente? *Democrația* cere ca toată lumea să posede *aceeași* abilitate de a modela procesul politic? Ce înseamnă o distribuire *corectă* a resurselor societății? Iată genul de întrebări la care teoreticienii politici încearcă să găsească un răspuns. Această disciplină are o istorie deopotrivă îndelungată și bogată. De la Platon și Aristotel din Academia din Grecia antică, până la Thomas Hobbes și Jean-Jaques Rousseau din tumultul Europei moderne timpurii, dar și titanii gândirii politice ai secolului al XX-lea precum Hannah Arendt și John Rawls, gânditorii politici occidentali au încercat să clarifice și să observe critic lumea din jurul lor, întrebându-se de ce era în felul acela și dacă nu cumva ar putea sau ar trebui să fie diferită.

Din câteva considerente, teoria politică este cât se poate de adecvată pentru analizarea interacțiunii dintre tehnologie și politică. În primul rând, canoanele gândirii politice conțin înțelepciunea care a supraviețuit civilizațiilor. Ea poate pune în lumină situații dificile viitoare și ne poate ajuta să ne dăm seama ce este în joc. Ar fi o nesăbuire din partea noastră să nu profităm de tezaurul ideilor pe care le avem la dispoziție, chiar dacă, în cele din urmă, vom decide că unele din ele trebuie actualizate sau „resetate". Teoriile politice ne oferă totodată *metode* de gândire despre lume, care ne ajută să ridicăm nivelul dezbaterilor deasupra aserțiunilor și a prejudecăților.

În opinia mea, lucrul cel mai bun în privința teoriilor politice este că acestea abordează temele și întrebările *majore* ale politicii. Ele oferă o vedere panoramică a pădurii politice,

acolo unde alte abordări s-ar putea pierde printre copaci (sau s-ar putea înțepeni printre ramuri). În cazul nostru, acest lucru este necesar pentru a respecta tema cărții de față. În situația în care considerăm că tehnologia ar putea avea un impact fundamental asupra condiției umane, atunci și analiza acestui impact ar trebui să fie una fundamentală. Iată de ce, în cartea de față, ne referim la patru concepte politice fundamentale:

Puterea: Cum îi domină cei puternici pe cei slabi
Libertatea: Ce este permis și ce este interzis
Democrația: Cum pot oamenii să guverneze
Justiția socială: Ce îndatoriri avem unii față de ceilalți

În perioade de mari schimbări, sugerez eu, merită să ne întoarcem la principiile inițiale și să ne gândim la aceste concepte în afara unui anumit regim juridic. Am reușit, astfel, să ne imaginăm un sistem superior celui pe care l-am moștenit.

Teoriile politice sunt utile și pentru că ne permit să reflectăm critic nu numai cu privire la politică, ci și despre cum *gândim* și *vorbim* despre aceasta. Conceptele sunt „gaura cheii prin care vedem și percepem inevitabil realitatea". Dacă doresc să-i vorbesc despre politică vecinei mele, nu sunt nevoit să încep de la zero. Știu că dacă afirm despre un anumit proces că este „nedemocratic", ea va ști destul de bine ce am vrut să spun și ce conotații am dorit să transmit, fără să fiu nevoit să-i explic ce este democrația și de ce ar trebui ca aceasta să fie considerată un lucru bun. Asta, pentru că facem parte din aceeași comunitate lingvistică și împărtășim o „rezervă comună de noțiuni" desprinse din istoria și mitologia noastră comună. Este un lucru convenabil.

Pe de altă parte, ce dorim să exprimăm despre politică poate fi uneori limitat de puținătatea cuvintelor de care dispunem. Unele lucruri par inexprimabile sau de neconceput pentru că rezerva comună de noțiuni încă nu s-a dezvoltat suficient pentru a le da glas. „Limitele limbajului meu", spunea Ludwig Wittgenstein, „sunt și limitele lumii mele".

În termeni politici, aceasta înseamnă că, chiar dacă am putea să vedem limpede viitorul, s-ar putea să nu dispunem de cuvintele adecvate pentru a-l descrie. Este motivul pentru care atât de des ne limităm viziunea asupra viitorului, la o versiune supraaccelerată a lumii în care trăim deja. „Dacă aș fi întrebat oamenii ce își doresc", spunea Henry Ford, primul producător de automobile în serie, „ei mi-ar fi răspuns că vor cai mai rapizi". Ford a recunoscut că poate fi greu să concepi un sistem radical diferit față de cel existent. Incapacitatea de a ne actualiza limbajul face acest lucru încă și mai dificil.

DE CE AM SCRIS ACEASTĂ CARTE

Prima oară când am dat cu adevărat de gustul teoriei politice a fost în facultate, unde m-am îndrăgostit de această disciplină sub îndrumarea atentă a unor profesori minunați. S-a trezit în mine o obsesie care nu m-a părăsit nici acum. (Recunosc, retrospectiv, că cea mai reușită poveste de iubire din timpul studenției mele se poate să fi fost o idilă pasională, dar condamnată la eșec, cu filosoful german G. W. F. Hegel.)

Lăsând pasiunea la o parte, ceva din disciplina „teoriilor politice" mă răscolea. Teoreticienii politici păreau să fie mândri de gândirea lor profundă asupra *istoriei* ideilor politice, însă, cu câteva excepții, erau aproape complet neinteresați de *viitorul* acestora. Mi s-a părut ciudat: de ce aceiași erudiți – atât de sensibili la context când scriau despre trecut – discutau despre politică de parcă lumea din anul 2050 avea să fie la fel ca aceea din 1950? Se părea că o mare parte a unor teorii politice foarte inteligente avea o aplicare practică redusă, pentru că nu se potrivea cu realitățile din prezent. Când am reflectat la politica din viitor, gândul m-a dus la Orwell, Huxley, Wells – toți, romancieri de la începutul secolului al XX-lea – mai degrabă decât la teoreticienii politici din secolul

al XXI-lea. Se pare că nu eram singurul: de la alegerea lui Donald Trump la președinția SUA, spre sfârșitul anului 2016, romanul lui Orwell, *1984*, a revenit pe lista cărților bestseller. Însă apare întrebarea: dacă vrem să înțelegem lumea așa cum va fi ea în anul 2050, chiar ar trebui să ne bazăm pe o lucrare de ficțiune scrisă în 1949?

După ce am absolvit facultatea și m-am implicat în propriile cauze politice modeste, supărătoarea mea stare de neliniște cu privire la faptul că teoria politică ar putea fi incapabilă sau refractară să abordeze provocările care se prefigurează ale generației mele a devenit o preocupare mai urgentă. Dacă tehnologia ar evolua atât de repede, încât ne-ar lipsi dotarea intelectuală să o mai înțelegem? Dacă, pe negândite, am fi pe punctul de a dezlănțui un viitor pe care nu l-am putea înțelege și, cu atât mai puțin, controla?

Doream niște răspunsuri, motiv pentru care am început să lucrez la cartea aceasta.

POLITICA TEHNOLOGIEI

Înainte de a trece mai departe, să începem cu o întrebare simplă: care este legătura dintre tehnologia digitală și politică?

Tehnologia în general

Noile tehnologii *permit* realizarea unor lucruri care mai înainte nu puteau fi înfăptuite, iar acestea *fac mai ușoară* înfăptuirea unor lucruri pe care le realizam deja. Aceasta este semnificația lor socială de bază. Cel mai adesea, noile oportunități create de tehnologie sunt de natură minoră: un mod ingenios de măcinare a boabelor de cafea, de exemplu, este puțin probabil să ducă la răsturnarea unui stat. Uneori, însă, consecințele pot fi importante. În cazul revoluției industriale, inventarea războaielor de țesut mecanice și a mașinilor de tricotat și de filat amenința să elimine locurile

de muncă ale pricepuților muncitori textiliști. Unii dintre aceștia, cunoscuți drept „ludiți", au pornit o revoltă violentă în regiunea rurală a Angliei, distrugând toate noile mașini pe care le întâlneau în drumul lor. Încă mai folosim termenul de „ludiți" pentru a-i descrie pe cei care opun rezistență la apariția unor tehnologii noi și perturbatoare.

Consecințele *economice* ale inovației, ca în cazul ludiților, vor impune adesea și o reacție politică. Însă noile tehnologii pot să aducă și provocări de natură *morală*. Peste câțiva ani, de exemplu, sistemele de realitate virtuală vor facilita comportamente care mai înainte erau de domeniul științifico-fantasticului, inclusiv (bunăoară) posibilitatea de a face sex virtual prin intermediul unor platforme capabile de stimulare a activității sexuale umane. Acest lucru ridică niște întrebări interesante. Ar trebui ca legea să permită să faci sex virtual cu un partener digital astfel conceput încât să arate ca o persoană din viața reală – o celebritate, de exemplu – fără cunoștința sau consimțământul persoanei respective? S-ar putea ca oamenii să manifeste sentimente puternice în această privință. Ce ar fi dacă o femeie ar face „sex" virtual cu încarnarea virtuală a prietenului celui mai bun al soțului ei? Ar părea greșit ca legea să permită așa ceva, însă ai putea susține că ceea ce faci în intimitatea lumii tale virtuale, fără să produci niciun rău nimănui, nu este treaba altcuiva și cu atât mai puțin a legii. Ca să dăm un exemplu și mai puternic: ce-ar fi să faci sex virtual cu avatarul unui copil, în condițiile în care nu este molestat niciun minor în cursul acelei experiențe?

Acestea sunt niște întrebări noi. Nu a mai trebuit să răspundem până acum la ele. Poate că v-ați format deja niște opinii în privința lor. Poate că acele opinii diferă de ale mele. Exact asta și este ideea: sunt niște întrebări inerent *politice* și ar trebui să li se răspundă cu atenție, cu referire la o serie acceptabilă de principii morale. Așa cum a remarcat un distins autor, noile tehnologii ne pot determina să ne reexaminăm

vederile politice, tot așa cum un preparat nou din meniul restaurantului preferat ne poate determina să ne reevaluăm preferințele culinare.

Unele tehnologii sunt „inerent politice" prin aceea că necesită, de fapt, „anumite tipuri de relații politice" sau măcar sunt compatibile cu acestea. Când Langdon Winner scria, prin anul 1980, cita drept exemplu puterea nucleară: dacă urmează să adoptați un sistem bazat pe energia nucleară, atunci trebuie să acceptați și faptul că va trebui ca o „elită tehnico-științifică-industrial-militară" puternică să-i administreze și să-i supravegheze instalațiile și infrastructura.

Mai există și alte tehnologii, care nu sunt inerent politice, dar care sunt politizate de contextul în care funcționează. În Marea Britanie, de exemplu, o armă de foc autorizată este, de regulă, o tehnologie utilizată la vânarea animalelor sălbatice. Armele nu fac parte din cultura dominantă, iar oamenii sunt, în general, foarte fericiți cu ideea unei reglementări stricte a acestora. Numai că, în SUA, dreptul de a purta arme este garantat de cel de-al doilea Amendament [al Constituției – n. tr.]. Opoziția culturală față de această reglementare este mult mai tranșantă. Aceeași tehnologie, un context politic diferit.

O ultimă legătură, mai subtilă, între tehnologie și politică este aceea că invențiile au modul lor de a se insinua în viața politică și intelectuală. Un bun exemplu în acest sens este ceasul mecanic. Așa cum arată Otto Mayr, unele civilizații antice și-au imaginat statul ca pe un corp uman, fiecare membru component al comunității politice alcătuind extremități precum o „mână" sau un „picior". În Renașterea târzie, metaforei corpului uman i s-au alăturat și altele, cum ar fi „corabia statului". După Copernic, monarhul a ajuns să fie văzut ca un mare soare în jurul căruia gravitau servitorii săi. Între secolele al XVI-lea și al XVIII-lea, metafora dominantă a fost aceea a ceasului, o mașinărie ingenioasă care inspira o „venerație fără precedent". După un timp, gânditorii au ajuns

să vadă politica din perspectiva mecanismului unui ceas. Armonia, cumpătarea și consecvența au devenit idealurile politice predominante. Oamenii tânjeau după un „ om de stat-inginer atotvigilent" care să poată să „prevadă, dar și să remedieze toate problemele posibile". Așa se face că o anumită tehnologie și o serie de valori politice au mers mână în mână, atingând apoteoza în secolul al XVII-lea, odată cu „extraordinara sa producție de ceasuri" și cu „înflorirea vădită a concepției autoritare despre ordine".

Și imaginația politică a secolului al XX-lea a fost nuanțată de tehnologie. Chiar și anterior erei digitale, perspectiva unor calculatoare atotputernice a inspirat o mare parte a artei și a literaturii de ficțiune. În *Mașina se oprește* (1928), E. M. Forster a zugrăvit o lume în care oamenii erau subordonați Mașinii, un sistem tehnologic global care supraveghea și controla fiecare aspect al existenței umane.

Unii scriitori au încercat să definească ceea ce considerau a fi ideologia timpurilor noastre. Evgheni Morozov, de exemplu, a scris despre „Doctrina Google" („încrederea entuziastă în forța eliberatoare a tehnologiei, însoțită de impulsul irezistibil de a înrola start-up-urile din Silicon Valley în lupta mondială pentru libertate"); „cyber-utopismul" („convingerea naivă în natura emancipatoare a comunicării on-line, care se bazează pe refuzul încăpățânat de a-i recunoaște dezavantajele") și „soluționismul" („reformularea tuturor situațiilor sociale complexe fie ca probleme clar definite cu soluții clare, estimabile, fie ca procese transparente și de la sine înțelese care pot fi optimizate cu ușurință – cu condiția aplicării algoritmilor potriviți").

Opinia mea personală este că tehnologiile în chestiune sunt de dată prea recentă pentru a ști ce efect de durată vor avea asupra gândirii noastre politice. A fost nevoie de sute de ani pentru ca ideea ceasului mecanic să se impregneze în viața politică și intelectuală a Europei. Cum alegem să interpretăm

tehnologiile vremurilor noastre și cum ne modelează acestea, la rândul lor, percepția asupra lumii sunt chestiuni care urmează încă a fi determinate.

Tehnologiile digitale

Cam atât, pentru moment, despre relația generală dintre tehnologie și politică. Această carte se referă, însă, la tehnologiile *digitale* – care mai înainte erau numite tehnologii ale informației și comunicațiilor sau TIC. Și se pare că acestea nu sunt numai inerent politice; sunt *hiper*politice. Asta, pentru că vizează două componente esențiale ale vieții politice: *comunicațiile* și *informația*.

Întreaga ordine politică se întemeiază pe coordonare, cooperare sau control. Este imposibil să organizezi viața colectivă fără cel puțin una din acestea trei. Și niciuna dintre ele nu este posibilă în lipsa unui sistem de schimb de informații fie între oamenii obișnuiți, fie între cei care conduc și cei care sunt conduși. De aceea este atât de important *limbajul*. După cum spune James Farr, fără limbaj, politica nu ar fi numai „imposibil de descris", ci și de-a dreptul „imposibilă":

> Statele emergente nu și-ar putea declara independența, liderii nu și-ar putea instrui adepții, cetățenii nu ar putea protesta împotriva războaielor, iar instanțele nu ar putea să condamne infractorii. Nici noi nu am putea să criticăm, pledăm, promitem, susținem, îndemnăm, cerem, negociem, să cădem la înțelegere, să facem compromisuri, să consiliem, informăm, chestionăm, sfătuim sau să consimțim. Să ne imaginăm politica fără aceste acțiuni ar însemna să ne imaginăm o politică de nerecunoscut.

Yuval Noah Harari observă, în mod elegant, că limbajul a jucat un rol esențial în perioada timpurie a politicii umane. Un număr redus de sunete și semne le-a permis strămoșilor noștri să emită un număr infinit de propoziții, fiecare având o anumită semnificație. Aceasta le-a oferit posibilitatea să

vorbească unii despre ceilalţi, să vorbească despre lucruri importante şi complicate din lumea din jurul lor, dar şi despre lucruri lipsite de orice consistenţă materială, precum miturile şi poveştile. Astfel de legende ţin comunităţile unite şi în zilele noastre.

Este posibil ca scrierea să fi apărut ca un mod de reflectare a realităţii, „însă a devenit, treptat, un mod puternic de reconfigurare a realităţii." *Domesday Book**, elaborată în anul 1086, a fost realizată la cererea lui William Cuceritorul pentru a se afla întinderea şi valoarea proprietăţilor oamenilor, în vederea impozitării mai eficiente a acestora. Conţinutul registrului era „bătut în cuie" şi nu putea fi contestat. În 1179, un comentator scria despre acest „sever şi cumplit ultim bilanţ", că:

> decizia sa nu poate fi anulată sau respinsă fără teama de consecinţe. De aceea am şi numit această carte „Judecata de Apoi" (...), pentru că hotărârile sale, ca şi cele de la Judecata de Apoi, sunt irevocabile.

Nu conta decât ceea ce era trecut în registru. Puţine lucruri s-au schimbat de atunci. După cum spune Harari: „Oricine a avut vreodată de-a face cu autorităţile fiscale, sistemul de educaţie sau cu orice alt sistem birocratic stufos ştie că adevărul abia dacă mai contează. Ce apare scris în formularul tău este cu mult mai important."

Limbajul, însă, nu este suficient. Şi comunităţile politice avansate trebuie să fie capabile să proceseze un volum mare de *informaţii* – de la rata sărăciei şi a creşterii economice, până la nivelurile fertilităţii, şomajului şi imigraţiei. Dacă vrei să guvernezi o naţiune, trebuie mai întâi să o cunoşti.

* *Domesday Book* („Cartea Judecăţii de Apoi") este un registru al anchetei privind terenurile şi animalele din cea mai mare parte a Angliei şi o parte din Ţara Galilor. În prezent se află la Arhivele Naţionale, Kew, Richmond upon Thames, în sud-vestul Londrei. (n. tr.)

Revoluționarii secolului al XVIII-lea, care au încercat să înlocuiască *ancien régimes** decadente, au înțeles bine acest lucru. Revoluția Franceză din 1789 a fost urmată de un efort intens de standardizare și raționalizare, de introducere a unui sistem unic de greutăți și măsuri, de împărțire a teritoriului în *départements*** și de promulgare a Codului Civil. Peste Atlantic, Părinții Fondatori ai Statelor Unite au prevăzut necesitatea unui recensământ decenal (o „Enumerare") chiar în primul articol al Constituției. Alexander Hamilton considera că guvernul federal ar trebui să constituie „centrul de informare", pentru că astfel ar fi cel mai în măsură să „înțeleagă amploarea și urgența pericolelor care amenință să-și facă apariția".

Legătura dintre informație și politică este fundamentală și și-a pus amprenta asupra vocabularului nostru. Cuvântul *statistică* provine de la termenul german din secolul al XVIII-lea, *Staatwissenschaft*, „știința statului", predată de profesorii universitari prințișorilor Sfântului Imperiu Roman. Într-adevăr, definiția funcțională a statisticii – „a face ca elemente anterior separate să existe împreună, oferind astfel realitate și consistență unor obiecte mai mari și mai complexe" – este aproximativ aceeași cu scopul politicii, cel dintâi, ca abstracție numerică, celălalt, ca realitate umană. Inițial, cuvântul englezesc *classified* [„clasificat" – n. tr.] avea un înțeles principal: sortarea informației pe taxonomii. În secolul al XIX-lea, pe când statul britanic devenea din ce în ce mai puternic, iar imperiul său câștiga în amploare, *classified* a ajuns să primească suplimentar și înțelesul său actual: informație deținută sub jurisdicția exclusivă a statului. Chiar și cuvântul *control* are o origine informațională, fiind derivat din verbul din latina Evului Mediu *contrarotulare*, care însemna să

* Vechile regimuri (în franceză, în original) (n. tr.)
** Departamente (în franceză, în original) (n. tr.)

compari ceva „cu rotulusurile", sulurile de hârtie care serveau drept înscrisuri oficiale.

În *Economie și societate* (1922), cea mai importantă lucrare sociologică de la începutul secolului al XX-lea, Max Weber anunța „instrumentul de precizie" al *birocrației* ca fiind modul cel mai avansat de organizare a activității umane de până atunci. Opinia lui Weber, justificată pe parcursul următoarelor câteva sute de ani, era că un plan unificat de organizare sistematică a informației, bazat pe „precizie, viteză, lipsă de ambiguitate, cunoașterea dosarelor, continuitate, discreție (...), unitate" ar constitui cel mai eficient sistem de control politic. Lecția aceasta nu a fost irosită de Franklin D. Roosevelt atunci când și-a asumat președinția Statelor Unite, în 1933. Celebrul său program de intervenție economică a fost asociat cu efortul, mai puțin seducător, dar nu mai puțin vital, de modificare a abordării statisticii de către guvernul federal. Perioada dintre anii 1935 și 1950 a fost martora apariției celor „trei instrumente principale ale statisticii sociale și economiei moderne": sondajele bazate pe eșantioane reprezentative, sistemul conturilor naționale și informatica. Dacă Hamilton, Weber și Roosevelt ar mai fi trăit încă, toți ar fi interesați de uriașele cantități de date pe care le producem în prezent și de sistemele tot mai performante pe care le utilizăm pentru procesarea lor.

Făcând un pas înapoi, putem acum să avansăm o ipoteză preliminară care să ne călăuzească pe parcursul cărții: aceea că modul în care culegem, stocăm, analizăm și comunicăm informația – în esență, modul în care o organizăm – este strâns legat de modul în care ne organizăm viața politică.[...]

Întorcându-ne din trecut spre viitor, trebuie să ne întrebăm cum vor afecta tehnologiile digitale transformaționale – tehnologii ale informației și comunicațiilor – sistemul *nostru* politic.

Aceasta este întrebarea care stă la baza cărții de față.

Ştim că efectele tehnologiei sunt diferite de la un loc la altul. Apariţia tehnologiei tiparului în China şi Coreea, de exemplu, nu a provocat acelaşi gen de schimbare care a urmat după apariţia tiparului lui Gutenberg în Europa, unde societatea era mai bine pregătită pentru o revoluţie religioasă şi politică. Asemenea diferenţe pot fi explicate, de regulă, prin prisma împrejurărilor economice şi politice. Cine deţine şi controlează o anumită tehnologie, modul în care este aceasta primită de public, dacă posibilele ei întrebuinţări sunt gândite dinainte şi dacă tehnologia este direcţionată spre un anumit scop, toate acestea îi vor influenţa impactul.

Asta înseamnă că ar trebui să nu ne grăbim să tragem concluzia că dezvoltarea unei anumite tehnologii va duce inevitabil sau implacabil la un anumit rezultat social. Gândiţi-vă la internet: pentru că structura reţelei sale era implicit potrivită pentru o organizaţie decentralizată şi neierarhizată, mulţi au prezis, cu toată încrederea, că existenţa on-line va fi destul de diferită de cea din lumea off-line. Însă lucrurile nu au prea evoluat în această direcţie. În principal ca o consecinţă a lumii comerciale şi politice în care şi-a făcut apariţia, internetul a intrat din ce în ce mai mult sub coordonarea şi controlul marilor companii şi al entităţilor politice care ne filtrează şi ne modelează experienţa on-line.

În plus, nu putem să considerăm că tehnologia este echivalentă cu progresul.[...]

Pe când Lenin descria comunismul drept „puterea Sovietelor plus electrificarea întregii ţări", Troţki şi-a dat seama că progresul tehnologic nu constituia un garant al progresului moral. „Alături de secolul al XX-lea", scria el, conviețuiesc şi „secolul al X-lea sau al XIII-lea":

> O sută de milioane de oameni folosesc electricitatea, dar continuă să creadă în puterea magică a semnelor prevestitoare şi a exorcismului. (...) Ce rezerve inepuizabile de întunecime

spirituală, ignoranță și sălbăticie posedă aceștia! (...) Tot ce ar fi trebuit eliminat din organismul națiunii în (...) cursul dezvoltării nestingherite a societății, ni se înfățișează astăzi, țâșnindu-i din gât.

Nu putem să luăm de bună niciuna dintre concluzii. Nici nu putem să considerăm că facultățile morale ni se vor dezvolta automat, odată cu invențiile noastre. Jocul încă nu s-a terminat.

CUPRINS

Cuvânt-înainte de Ciprian I. Tomoiagă 5
Notă despre titlu 11
Introducere 15

Capitolul 1 — *Puterea* 19
Capitolul 2 — *Datele personale informatizate* 43
Capitolul 3 — *Justiția* 75
Capitolul 4 — *Medicina* 110
Capitolul 5 — *Automobilele autonome* 150
Capitolul 6 — *Infracționalitatea* 185
Capitolul 7 — *Arta* 225

Încheiere 251
Note 259
Mulțumiri 313
Glosar 315
Index 321

Pentru comenzi și informații, contactați:
GRUPUL EDITORIAL CORINT
Departamentul de Vânzări
Str. Mihai Eminescu nr. 54A, sector 1, București, cod poștal 010517
Tel./Fax: 021.319.47.97; 021.319.48.20
Depozit
Calea Plevnei nr. 145, sector 6, București, cod poștal 060012
Tel.: 021.310.15.30
E-mail: vanzari@edituracorint.ro
Magazin virtual: www.edituracorint.ro

Format: 16/54x84;
Coli tipo: 22

Tiparul executat la:

Hannah Fry
(n. 1984) este matematiciană, conferențiar universitar doctor și autoare a unor celebre cărți de popularizare a științei, în care a tratat aspecte legate de aplicarea matematicii pentru rezolvarea problemelor lumii contemporane, precum și diversele chestiuni privind utilizarea algoritmilor, a programelor de calculator: *The Mathematics of Love: Patterns, Proofs and the Search for the Ultimate Equation* (2015), *The Indisputable Existence of Santa Claus* (2016).

Finalistă a premiului pentru carte științifică acordat de Royal Society

„O descriere onestă, scrupuloasă, profundă și plină de stil a ceea ce programele și algoritmii care au ajuns să ne conducă viețile pot și nu pot să facă. Un ghid extrem de accesibil, care trece cu ușurință prin domenii dintre cele mai diverse, fără a-l scuti însă pe cititor de obligația de a-și pune unele dintre cele mai grele și mai serioase întrebări despre societatea prezentului, dar și a viitorului... O carte care își merită locul în topurile de bestselleruri."

Oliver Moody, *The Times*

„Genial de clară! Hannah Fry creionează cele mai importante probleme de natură etică pe care le ridică folosirea algoritmilor în zilele noastre."

James McConnaiche, *Sunday Times*

Partener media

Carte recomandată de

ISBN: 978-606-793-598-1

www.edituracorint.ro